POLITECNICO DI TORINO
Repository ISTITUZIONALE

Optimum Receiver Design for MIMO Fading Channels

Original
Optimum Receiver Design for MIMO Fading Channels / Coluccia, Giulio. - (2008). [10.6092/polito/porto/2502249]

Availability:
This version is available at: 11583/2502249 since:

Publisher:
Politecnico di Torino

Published
DOI:10.6092/polito/porto/2502249

Terms of use:
Altro tipo di accesso

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

05 May 2023

POLITECNICO DI TORINO

SCUOLA DI DOTTORATO
Dottorato in
Ingegneria Elettronica e delle Comunicazioni – XXI ciclo

Tesi di Dottorato

Optimum Receiver Design for MIMO Fading Channels

Giulio Coluccia
mat. 136164

Tutore
Prof. Giorgio Taricco

Coordinatore del corso di dottorato
Prof. Ivo Montrosset

Gennaio 2009

Summary

This thesis describes the analytical design and the performance analysis of optimum receivers for Multiple Input – Multiple Output (MIMO) fading channels.

In particular, a novel *optimum receiver* for separately–correlated MIMO channels is proposed. This novel pilot–aided receiver is able to process jointly the pilot symbols, transmitted within each time frame as a preamble, and the information symbols and to decode the transmitted data in a single step, avoiding the explicit estimation of the channel matrix.

The optimum receiver is designed for the following two scenarios, corresponding to different transmission schemes and channel models:

1. Narrowband Rician fading MIMO channel with spatial separate correlation;
2. MIMO–OFDM Rician fading channel with space and frequency separate correlation.

For each system the performance of the optimum receiver is studied in detail under different channel conditions. The optimum receiver is compared with:

- the ideal *genie receiver*, knowing perfectly the Channel State Information (CSI) at no cost;
- the standard *mismatched receiver*, estimating the CSI in a first step, then using this imperfect estimate in the ideal channel metric.

Since the optimum receiver requires the knowledge of the channel parameters for the decoding process, an estimation algorithm is proposed and tested.

Moreover, a complexity analysis is carried out and methods for complexity reduction are proposed.

Furthermore, the narrowband receiver is tested in realistic conditions using measured channel samples.

Finally, a blind version of the receiver is proposed.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Multiple Input – Multiple Output (MIMO) communications have attracted the interest of many researchers during the last decade because of the promise of multiplying the achievable rate by a factor equal to the minimum number of antennas employed at the transmitter and at the receiver [1, 2, 3] at an affordable cost. This feature spurred researchers to develop channel codes properly designed for MIMO channels, the Space-Time Codes (STCs) [4, 5, 6].

However, most of the early works relied on assumptions that turned out to be critical to the actual achievement of the increased capacity:

1. channel state information (CSI) must be perfectly known at the receiver;

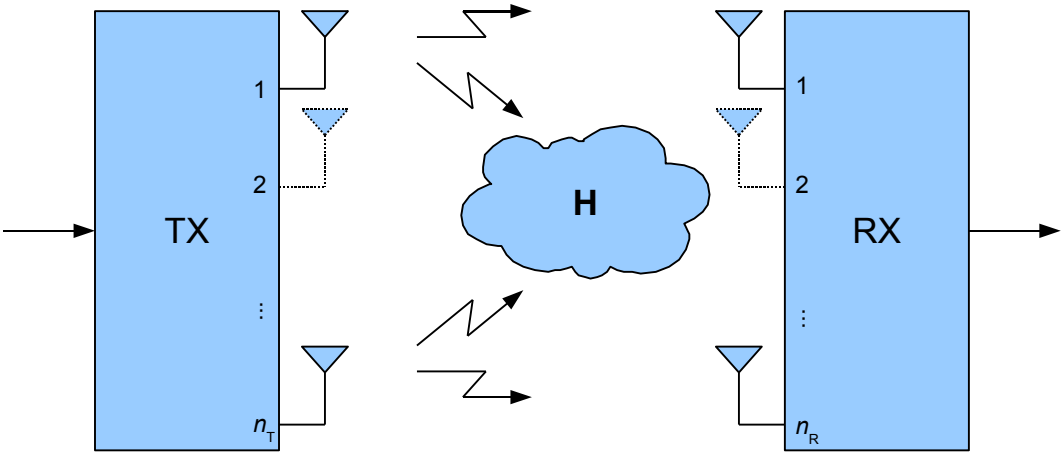


Figure 1.1. Transmission and reception with multiple antennas. The gains of each propagation path are described by the $n_R \times n_T$ matrix $\mathbf{H}$

2. the signal propagation paths must be uncorrelated;

3. there is no direct-path propagation, i.e., the signal arrives at the receiver only through scattering and reflections.

Only more recently, researchers realized the importance of those issues. The effect of path correlation was studied, among the others, in [7, 8, 9, 10, 11, 12]. The presence of a direct propagation path was taken into account in [13,14,15]. Methods for exploiting direct-path propagation and path correlation were proposed in [16,17, 18].

Though the general correlation of the propagation path gains is quite difficult to take into account in a MIMO system, many works have suggested the application of the separately correlated MIMO channel model [19, 9, 7, 8, 16], which was proposed in [20]. According to this model, path correlation is determined as the product of a receive-side and a transmit-side component. This allows one to write the channel matrix in a simple format based on two constant correlation matrices plus an inner matrix of independent and identically distributed circularly-symmetric complex Gaussian random variables. On one hand, this model has been blamed of oversimplifying the actual covariance structure of the MIMO channel [21]; on the other hand, a study supporting the use of the Kronecker model has been published by Maharaj, Linde, and Wallace [22]. They have shown that the maximum entropy MIMO channel model for known transmit and receive covariance matrices is actually different from the Kronecker model. However, the eigenvectors of the full covariance of the maximum entropy MIMO model are equal to those of the Kronecker model. For this reason, our tractation adopts the Kronecker model for the MIMO channel to make the development more manageable.

Other MIMO channel models have been proposed in the literature, such as Müller's random matrix model [23] based on the superposition of scattering components with different amplitudes and uniform phases; Sayeed's model [24] based on *virtual channel representation*; Tulino *et al.*'s Rician model [25] based on the decomposition of the channel matrix into the sum of a constant component plus the product of a left-hand unitary matrix by a complex Gaussian matrix with independent entries of different variances, by a right-hand unitary matrix; Weichselberger *et al.*'s Rayleigh model [26] based on a refined correlation structure which describes the average coupling between the eigenmodes of the two link ends.

In the first part of this thesis we focus on a narrowband separately correlated MIMO channel with imperfect CSI estimation obtained by inserting pilot symbols among the transmitted data.

In this setting, CSI is equivalent to the knowledge of the channel matrix $\mathbf{H}$ which contains the signal gains between all pairs of transmit and receive antennas. In this context we design an optimum transceiver structure that outperforms significantly other designs not exploiting the statistical properties of the MIMO channel.

It is well known that channel state information at the receiver (CSIR) is a fundamental requirement to communication systems. Essentially, digital receivers use it to remove the effect of passing the signal through the communication channel and implementing receivers without CSIR would always entail a significant error performance loss.

The recovery and estimation of sufficiently precise CSIR require a sizeable amount of resources. Nevertheless, receiver design assumes in many cases that perfect CSIR is available at no cost.

Our approach stems from the concepts developed by Cavers in [27] and Cavers and Ho [28] for single-input single-output (SISO) channels. Most digital communication receivers are designed under the assumption that CSIR is recovered by using pilot symbol insertion. This technique consists in periodically inserting *pilot symbols* in the data frame which are known both to the transmitter and to the receiver. The presence of channel noise implies that the recovered CSIR is affected by estimation errors.

In order to reduce the effect of CSIR inaccuracy, a considerable fraction of the available transmission power has to be spent for transmitting pilot signals. This reduces either the power and the throughput efficiencies because only part of the available resources are actually used to transmit data.

The impact of CSIR estimation accuracy further increases in the case of MIMO communication systems, for the obvious reason that the received signal depends on a large number of channel gains.

Many works have addressed the issue of CSIR recovery for MIMO systems in the technical literature. They can be divided into two main categories:

1. those addressing the issue of channel capacity and

2. those addressing the issue of error performance.

Among the former, it is worth citing Marzetta [29], who studied the MIMO channels with BLAST and orthogonal training signals system and showed that maximizing the overall transmission rate requires the transmission interval to be used half for training and half for data transmission. It was also shown that the training-sequence length should be approximately proportional to the number of transmit antennas. More recently, Hassibi and Hochwald [30] derived a lower bound to the channel capacity with imperfect CSIR derived from pilot symbol insertion. Their lower bound consists in merging the channel estimation error and noise into a single disturbance component which is assumed to be iid and Gaussian. In a related area, Yoo and Goldsmith [31] investigated the effect of imperfect CSIR and CSIT (channel state information at the transmitter) in the case of broadcast channels. Information-theoretic analyses aimed at finding the MIMO channel capacity with

imperfect CSI and pilot-aided channel estimation can be found in [30, 32]. More recent information-theoretic analyses address the case of separately-correlated Rician fading MIMO channels [33, 34, 35].

Among the latter (papers addressing the issue of error performance), Taricco and Biglieri [36] proposed two types of optimum receivers (according to different optimality criteria) and compared their error performance with independent narrowband Rayleigh fading and space–time trellis coding (STTC).

Here, we focus on the latter topic, i.e. the development of optimum detection algorithms aiming at minimizing the error probability after decoding (still using pilot-aided channel estimation) and apply them to selected examples of trellis STCs. Our approach is inspired by Tarokh *et al.* [37] who examine STCs in the presence of channel-estimation errors, though their analysis is partly affected by a flaw [38]. Following the methods proposed by Taricco and Biglieri [36] for the independent Rayleigh MIMO channel, we consider two receiver structures, hereafter referred to as *mismatched receiver* and *optimum receiver*.

The *mismatched receiver* estimates the channel matrix by using pilot symbols with maximum-likelihood (ML) or minimum mean-square error (MMSE) estimation of the MIMO channel matrix, and then decodes the transmitted code word by using the previous estimate as if it were exact. The problem of optimum parameter estimation with additive Gaussian noise received considerable attention in the literature (see, e.g., [39]).

The *optimum receiver* processes jointly the received signal and the pilot symbols to detect the information symbols. This receiver maximizes the *a-posteriori* probability of the data symbols given the pilot symbols and all the received signal samples (pilots and data, as suggested in [30]). A different approach, addressing V-BLAST, was recently proposed by Lee and Chun [40], who investigated a modified V-BLAST detection algorithm to reduce the effects of CSIR inaccuracy.

The principles behind the optimum receiver design are very intuitive but its derivation leads to complex decision metrics whose calculation is computationally intense and has not been considered earlier in the technical literature for the separately-correlated Rician MIMO channel.

However, complexity can substantially be reduced by means of iterative calculation of the decision metric, as we show in the paper. The resulting decoding scheme applies neatly to the case of trellis STCs.

Besides the higher complexity, the optimum receiver requires the estimation of statistic parameters describing the channel matrix. However, we show that a simple parameter estimation algorithm provides excellent results in terms of achievable frame error rate (FER).

Moreover, we apply the concepts developed to a set of measured channel realizations in order to assess how the proposed receiver performs in a real scenario.

Among the available channel measurements (obtained by *ftw.* [41]), we consider

two scenarios: the *urban* scenario (outdoor, with a moving transmitter and fixed receiver) and the *rich scattering* scenario (indoor, with fixed transmitter and receiver). Since the measurements were based on a high number of transmit/receive antennas, we obtain by properly scheduling the sample extraction different channel correlation scenarios depending on the corresponding antenna spacing.

Although the separately correlated Rician fading model is known to simplify the actual covariance structure of the MIMO channel [21], it is shown among the results proposed in this thesis that receivers based on it achieve good error performance in simulations based on *measured* narrowband channels in real environments.

However, it must be considered that pilot symbol insertion entails a rate loss that may be considerable in some cases and for particular applications, such as for low–complexity devices.

An alternative approach avoiding this rate loss is based on the use of *blind* detection techniques. Several blind detection schemes have been proposed in the literature for MIMO applications. Unitary space–time codes were developed for the independent Rayleigh channel in [42]. More recently, orthogonal space–time codes have been studied in [43] and a sequential Monte-Carlo technique is proposed in [44] to perform blind detection.

In this work a new blind detection scheme is derived for the narrowband separately correlated Rician fading MIMO channel by extending the approach explained previously.

This blind receiver is based on the perfect knowledge of channel distribution information at the receiver (CDIR), which can be obtained during a very short initial training phase, and it reaches the maximum achievable throughput of the channel.

Our approach bears some similarity to the Generalized Likelihood Radio Test (GLRT) receiver proposed in [45, 46] for single-input single-output (SISO) channels and in [47, 48] for MIMO channels. Consistently with [47], we show that the GLRT receiver metric converges to that of our blind receiver as the SNR grows asymptotically large, provided that the code words have full row rank.

In the second part of the thesis, we extend the approach to a MIMO–OFDM system which follows closely the assumptions introduced by the 802.11n draft 2.0 standard [49] as far as concerns the data and pilot frame structure.

The wideband block fading MIMO–OFDM channel matrix model is assumed to be completely specified by its first– and second–order statistics. This channel model allows for spatial correlation at the transmitter and receiver side, as well as for spectral correlation between OFDM subcarriers.

The Kronecker assumption is critically assessed by numerical results as far as concerns the error performance achieved by the optimum receiver (with imperfect CSIR knowledge) against the ML genie receiver (with perfect CSIR knowledge). We show that the optimum receiver, even in its suboptimum reduced–complexity

version with imperfect channel parameter knowledge, performs very close to the genie receiver in the considered cases, whereas the standard mismatched receiver loses several dBs.

1.1 Thesis outline and previously published papers

This thesis is composed by two main chapters.

After the introductory chapter 1, in chapter 2 the optimum receiver for narrowband MIMO channel is presented.

First, the system model and the the MIMO channel model are introduced in section 2.1. The transmission scheme and the frame structure are defined, and the details regarding the channel matrix are explained.

Section 2.2 illustrates the receiver architectures considered in this paper, namely the ML or MMSE mismatched receivers and the optimum receiver.

The maximum-likelihood metric is derived in section 2.2.3 and an iterative algorithm for the optimum receiver which is suitable to Viterbi decoding of trellis STCs is provided in section 2.2.4. This algorithm extends, in a nontrivial way, earlier results presented in [36] for the uncorrelated Rayleigh fading MIMO channel. The iterative formulation of the metric is given in order to reduce the computational complexity, that is studied in detail in section 2.2.5.

At the end of this section the problem of estimating the parameters of the channel matrix distribution according to the separately-correlated model assumed is considered. In section 2.2.6 a simple estimation algorithm that is subsequently analyzed by simulation is provided.

Section 2.3 collects numerical results for several illustrative examples. The nonzero mean uncorrelated MIMO channel is studied to assess the effect of the presence of a direct path on the FER performance. Both zero-mean and nonzero mean correlated MIMO channels are analyzed with the optimum receiver in order to establish the impact of correlation on the FER performance. The mismatched and optimum receivers are then compared to each other in the case of zero-mean correlated fading. The relative complexity of these receivers is discussed under the assumption of trellis STC decoding. Finally, simulation results are provided, showing that a simple parameter estimation algorithm (based on pilot symbol processing) allows to approach very closely the performance of the optimum receiver (based on ideal knowledge of the statistic parameters).

The receiver principle and some preliminary results were presented in the *44-th Annual Allerton Conference On Communication, Control, and Computing*, Monticello, Illinois, USA, September 2006, and published in [50]. A deep analysis of the receiver design and of the performance in a simulated environment was published in [51].

Section 2.4 describes the performance of the optimum receiver in a measured environment, i.e. using channel samples measured on the field. These results were presented in the *IEEE 8th Workshop on Signal Processing Advances in Wireless*

Communications, Helsinki, Finland, June 2007 and published in [52].

A blind version of the optimum receiver is presented in section 2.5. Its design and performance can be found also in [53].

Chapter 3 contains the design and the performance analysis of the optimum receiver for MIMO–OFDM channel.

The MIMO-OFDM system is described in section 3.1, where the channel matrix correlation model is proposed. Then, the mismatched (section 3.2.2) and optimum (section 3.2.3) receivers are designed. An iterative formulation of the optimum receiver is described in section 3.2.4, followed by an analysis of the complexity (section 3.2.5). A complexity reduction strategy based on eigenvalues decomposition is proposed in section 3.2.6. At last, numerical results are analyzed in section 3.3.

Again, a first analysis and performance overview was presented in *IEEE Global Communications Conference*, New Orleans, LA, USA, December 2008, while a complete analysis of the receiver design, especially from the point of view of complexity issues, is contained in [54].

Chapter 4 contains the concluding remarks.

1.2 Notation and Definitions

The analysis carried out in this thesis uses a huge amount of mathematical notation related to matrix algebra. In this section a brief overview of the main definitions can be found. Refer to [55] for further details.

Column vectors are denoted by lowercase boldface characters. The nth element of a vector $\mathbf{x}$ is $(\mathbf{x})_n$.

Matrices are denoted by uppercase boldface characters. The (m,n)th element of a matrix $\mathbf{A}$ is $(\mathbf{A})_{mn}$. The nth row of a matrix $\mathbf{A}$ is $(\mathbf{A})_n$.

The transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^{\mathsf{T}}$. The Hermitian transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^{\mathsf{H}}$.

The trace of a matrix is $\mathrm{Tr}\,(\mathbf{A}) = \sum_i (\mathbf{A})_{ii}$. The exponential of the trace of a matrix is $\mathrm{etr}\,(\mathbf{A}) = \exp(\mathrm{Tr}\,(\mathbf{A}))$.

The Frobenius norm of a matrix $\mathbf{A}$ is $\|\mathbf{A}\|$; its square can be written as $\|\mathbf{A}\|^2 = \mathrm{Tr}\left(\mathbf{A}\mathbf{A}^{\mathsf{H}}\right)$. The spectral norm of a matrix $\mathbf{A}$ is $|||\mathbf{A}|||$ and it is defined as $|||\mathbf{A}||| = \max_i \lambda_i$, where λ_i is the set of eigenvalues of the matrix $\mathbf{A}^{\mathsf{H}}\mathbf{A}$.

The notation $\mathbf{A} \otimes \mathbf{B} = [(\mathbf{A})_{ij}\mathbf{B}]$ (written in block matrix form, where the pair i,j spans the range of indexes of $\mathbf{A}$) denotes the Kronecker product of $\mathbf{A}$ times $\mathbf{B}$; the Kronecker product is endowed with several important properties:

1. $\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C}$;
2. $\prod_k (\mathbf{A}_k \otimes \mathbf{B}_k) = (\prod_k \mathbf{A}_k) \otimes (\prod_k \mathbf{B}_k)$;
3. $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$;
4. $\det(\mathbf{A} \otimes \mathbf{B}) = \det(\mathbf{A})^n \det(\mathbf{B})^m$ if $\mathbf{A} \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$.

The notation $\mathbf{A} \oplus \mathbf{B} = \mathrm{diag}\big(\mathbf{A},\mathbf{B}\big)$ denotes the matrix direct sum [55] of $\mathbf{A}$ and $\mathbf{B}$. The following property holds: given two (or more) square matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathrm{Tr}\,(\mathbf{A} \oplus \mathbf{B}) = \mathrm{Tr}\,(\mathbf{A}) + \mathrm{Tr}\,(\mathbf{B})$.

The notation $\mathrm{vec}\,\{\mathbf{A}\}$ denotes the column vector obtained by stacking the columns of $\mathbf{A}$ on top of each other from left to right.

The matrix $\mathbf{A}^{1/2}$ denotes the unique positive semidefinite matrix square-root of a positive semidefinite matrix $\mathbf{A}$ and is defined as $\mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}^{\mathsf{H}}$ if $\mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathsf{H}}$ is the singular-value decomposition (SVD) [55] of $\mathbf{A}$. If $\mathbf{A}$ is positive definite then $\mathbf{A}^{-1/2}$ is defined by $\mathbf{A}^{-1/2} = (\mathbf{A}^{-1})^{1/2}$.

The notation $\mathbb{E}_{\mathbf{H}}\,[(\cdot)]$ represents the expectation of $(\cdot)$ with respect to the random matrix $\mathbf{H}$.

The notation $\mathbf{x} \sim \mathcal{N}_c(\bar{\mathbf{x}},\mathbf{R}_x)$ means that the complex random column vector $\mathbf{x}$ is circularly-symmetric Gaussian distributed, its mean is $\bar{\mathbf{x}} = \mathbb{E}\,[\mathbf{x}]$, its covariance matrix is $\mathbf{R}_x = \mathbb{E}\left[(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^{\mathsf{H}}\right]$, and its probability density function is given by

$$p\,(\mathbf{x}) = \det(\pi\mathbf{R}_x)^{-1} \exp\left(-(\mathbf{x}-\bar{\mathbf{x}})^{\mathsf{H}}\mathbf{R}_x^{-1}(\mathbf{x}-\bar{\mathbf{x}})\right) \ .$$

1.3 List of Acronyms

As usual in technical literature, a list of the acronyms that can be found throughout the text is provided here.

ARQ: Automatic Repeat–reQuest

BR: Blind Receiver

CDI: Channel Distribution Information

CDIR: Channel Distribution Information at the Receiver

CDIT: Channel Distribution Information at the Transmitter

CSI: Channel State Information

CSIR: Channel State Information at the Receiver

CSIT: Channel State Information at the Transmitter

CSZMCG: Circularly–Symmetric Zero Mean Complex Gaussian

DFT: Digital Fourier Transform

FER: Frame Error Rate

GLRT: Generalized Likelihood Ratio Test

GR: Genie Receiver

iid: Independent and Identically Distributed

IDFT: Inverse Digital Fourier Transform

LOS: Line Of Sight

MCS: Modulation and Coding Scheme

MIMO: Multiple Input – Multiple Output

ML: Maximum Likelihood

MMSE: Minimum Mean Square Error

MR: Mismatched Receiver

OFDM: Orthogonal Frequency Division Multiplexing

OR: Optimum Receiver

pdf: Probability Density Function

PR: Pilot–aided Receiver

SISO: Single Input – Single Output

SNR: Signal–to–Noise Ratio

STC: Space–Time Code

STBC: Space–Time Block Code

STTC: Space–Time Trellis Code

SVD: Singular Value Decomposition

TR: Trained Receiver

V-BLAST: Vertical Bell Labs LAyered Space–Time

Chapter 2

Optimum narrowband receiver

In this chapter, two receiver structures based on pilot symbol-aided channel estimation are considered for the separately-correlated Rician fading MIMO channel.

1. A *mismatched receiver*, which decodes the received signal by first using maximum likelihood (ML) or minimum mean-square error (MMSE) estimation of the MIMO channel matrix, and then by assuming that the estimate is exact.

2. An *optimum receiver*, which does not estimate explicitly the channel matrix but jointly processes the received pilot and data samples assuming known channel distribution.

The main focus is on the optimum receiver. First, the optimum detection algorithm for the separately-correlated Rician fading MIMO channel is derived. Then, an iterative implementation suitable for trellis space–time decoding is proposed in order to reduce the algorithm complexity.

Numerical results are presented for a 2×2 MIMO system with Rayleigh/Rice correlated/uncorrelated fading and a simple trellis space–time code. These results show that a substantial gain is available with the optimum receiver either in terms of E_b/N_0 and system throughput (in both cases accounting for pilot-symbol rate reduction). Finally, the effect of parameter estimation (required by the optimum receiver) is studied and it is shown that this effect is almost unnoticeable in the case considered.

Moving on, the performance of the optimum pilot-aided receiver scheme, designed for the separately-correlated Rician fading MIMO channel, is analyzed over realistic MIMO scenarios obtained by channel measurements. The channel parameters required by the receiver are estimated during the data transmission without any throughput loss. Simulation results show that the optimum receiver performance is very close to the performance of a receiver with ideal channel estimation.

In the end of the chapter, a blind receiver scheme for narrowband separately-correlated Rician block fading coded MIMO systems is described. It is shown that this receiver (which can be interpreted as a limiting case of the optimum pilot-aided receiver) attains the maximum achievable throughput at sufficiently high E_b/N_0 ratio. The influence of the E_b/N_0 ratio and of the channel Rice factor on the receiver performance are studied and it is shown that, as either of these parameters increases, the advantage of the blind versus the pilot-aided receiver becomes more sensible. The results obtained support the intuition that, if the line-of-sight component is sufficiently strong, a blind scheme using only the knowledge of the fading statistics parameters matches or outperforms pilot-aided detection.

2.1 System Model

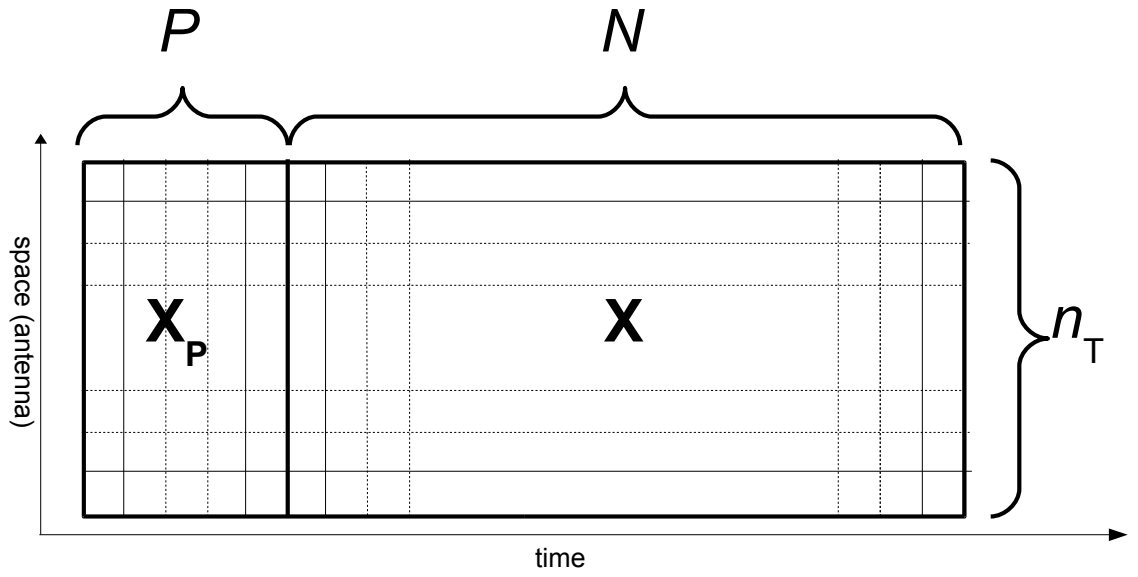


Figure 2.1. The frame structure

Consider a single user narrowband MIMO channel with n_T transmit antennas and n_R receive antennas. It is assumed that the fading process is sufficiently slow to consider it constant for the duration of a time frame (*quasi-static* assumption). A time frame (see Figure 2.1) is composed of P pilot symbol intervals and N data symbol intervals, both of duration equal to T time units. Thus, during one time frame, the transmitter sends a $n_\mathsf{T} \times P$ matrix $\mathbf{X}_\mathsf{P}$ of pilot symbols (the (i,n)th component of $\mathbf{X}_\mathsf{P}$ is transmitted from the ith antenna at the nth symbol interval in the time frame) followed by a $n_\mathsf{T} \times N$ matrix $\mathbf{X}$ of data symbols (the (i,n)th component of $\mathbf{X}$ is transmitted from the ith antenna at the $(P+n)$th symbol interval in the time frame).

It is assumed that pilot symbols and data symbols have average energy E_P and E_S, respectively. Therefore,

$$E_\mathsf{P} = \frac{1}{n_\mathsf{T} P}\mathbb{E}\left[\|\mathbf{X}_\mathsf{P}\|^2\right] = \frac{1}{n_\mathsf{T} P}\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}\right)\right]$$

and

$$E_\mathsf{S} = \frac{1}{n_\mathsf{T} N}\mathbb{E}\left[\|\mathbf{X}\|^2\right] = \frac{1}{n_\mathsf{T} N}\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{X}\mathbf{X}^\mathsf{H}\right)\right]$$

The $n_\mathsf{T} \times P$ *pilot symbol matrix* $\mathbf{X}_\mathsf{P}$ is used by the receiver to recover the unknown CSI.

As quasi-static fading is considered, the receiver observes the following sample matrices:

$$\mathbf{Y}_{\mathsf{P}} = \mathbf{H}\mathbf{X}_{\mathsf{P}} + \mathbf{Z}_{\mathsf{P}} \qquad \text{and} \qquad \mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{2.1}$$

Here,

$$\mathbf{H} = \begin{pmatrix} (\mathbf{H})_{11} & \cdots & (\mathbf{H})_{1n_{\mathsf{T}}} \\ \vdots & \ddots & \vdots \\ (\mathbf{H})_{n_{\mathsf{R}}1} & \cdots & (\mathbf{H})_{n_{\mathsf{R}}n_{\mathsf{T}}} \end{pmatrix}$$

is an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ *channel matrix* whose characteristics will be discussed in Section 2.1.1. $\mathbf{Z}_{\mathsf{P}}$ and $\mathbf{Z}$ are two $n_{\mathsf{R}} \times P$ and $n_{\mathsf{R}} \times N$, respectively, matrices representing the receiver noise with independent circularly-symmetric zero-mean complex Gaussian entries of variance N_0. The pilot symbol matrix $\mathbf{X}_{\mathsf{P}}$ is assumed to be a full-rank $n_{\mathsf{T}} \times P$ matrix with $P \geq n_{\mathsf{T}}$ (see Section 2.2.2 for further details).

2.1.1 Channel Matrix

The (i,j)th entry $(\mathbf{H})_{ij}$ of the channel matrix $\mathbf{H}$ represents the gain of the signal arriving from the jth transmit to the ith receive antenna. It is assumed that the fading channel is Rician and separately correlated [19, 9, 7, 8] so that the channel matrix can be written as

$$\mathbf{H} = \bar{\mathbf{H}}_0 + \mathbf{R}^{1/2}\bar{\mathbf{W}}\mathbf{T}^{1/2} \tag{2.2}$$

where

$$\bar{\mathbf{H}}_0 = \sqrt{\frac{K}{K+1}}\mathbf{H}_0 \qquad \text{and} \qquad \bar{\mathbf{W}} = \sqrt{\frac{1}{K+1}}\mathbf{W}$$

The $n_{\mathsf{R}} \times n_{\mathsf{R}}$ matrix $\mathbf{R}$ and the $n_{\mathsf{T}} \times n_{\mathsf{T}}$ matrix $\mathbf{T}$ are the receive and transmit correlation matrices, respectively. Both matrices are assumed to be positive definite. Their elements represent the correlation between antenna pairs at the receiver and transmitter. The $n_{\mathsf{R}} \times n_{\mathsf{T}}$ matrix $\mathbf{H}_0$ is normalized in order to have squared Frobenius norm $\|\mathbf{H}_0\|^2 = n_{\mathsf{T}} n_{\mathsf{R}}$. The $n_{\mathsf{R}} \times n_{\mathsf{T}}$ matrix $\mathbf{W}$ has independent circularly-symmetric zero-mean complex Gaussian entries with unit variance. Here, K is the Rice factor, which represents the relative strength of the line-of-sight component. Following the definition in [56], it can be seen that

$$K = \frac{\|\bar{\mathbf{H}}_0\|^2}{\mathbb{E}\left[\|\mathbf{H} - \bar{\mathbf{H}}_0\|^2\right]}$$

Thus, independently of K the following relation holds,

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)\right] = n_{\mathsf{T}} n_{\mathsf{R}} \ . \tag{2.3}$$

It can be seen that the covariance of the entries of $\mathbf{H}$ is given in terms of a product of the entries of the correlation matrices $\mathbf{R}$ and $\mathbf{T}$. Namely

$$\operatorname{cov}\left((\mathbf{H})_{ij},(\mathbf{H})_{i'j'}\right) = \frac{1}{K+1}(\mathbf{R})_{ii'}(\mathbf{T})^*_{jj'} \tag{2.4}$$

holds, where $\operatorname{cov}(X,Y) \triangleq \mathbb{E}[XY^*] - \mathbb{E}[X]\,\mathbb{E}[Y]^*$. The equivalence between (2.2) and (2.4) is stated in the following lemma.

Lemma 2.1.1. The definition of (2.2) satisfies the covariance condition (2.4).

Proof. See Appendix A.1. □

2.1.2 Estimation of the channel parameters

Fading variations occur over different time scales: short term variations are captured by the channel matrix $\mathbf{W}$, which changes from one frame to another; long term variations are accounted for by the channel matrix parameters $\mathbf{H}_0$, $\mathbf{R}$, and $\mathbf{T}$, which are supposed to remain constant for a sufficiently large number of time frames [18].

In the following the ideal case when the channel parameters are exactly known at the receiver and the more realistic case of a receiver which estimates them by using sample averages during a preliminary training phase (Section 2.2.6) are considered.

2.2 Receiver architectures

Two types of receiver architectures are considered:

1. the mismatched receiver, which estimates the channel matrix by using the pilot symbols and uses this estimate as if it were exact;

2. the optimum receiver, which processes jointly the received signal and the pilot symbols to decode the transmitted code word.

It is also considered as a benchmark the genie receiver, i.e. an ideal receiver having perfect knowledge of channel coefficients, provided by some genie device at no cost.

2.2.1 Genie Receiver

The ideal genie-aided receiver (yielding a lower bound to the FER) output is given by:

$$\widehat{\mathbf{X}}_{\mathsf{GR}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{GR}}(\mathbf{X}|\mathbf{Y}) \ , \tag{2.5}$$

where

$$\mu_{\mathsf{GR}}(\mathbf{X}) = \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \tag{2.6}$$

2.2.2 Mismatched receiver

This receiver operates in two stages. In the first stage, it estimates $\mathbf{H}$ by using $\mathbf{X}_{\mathsf{P}}$ and $\mathbf{Y}_{\mathsf{P}}$ according to an ML criterion:

$$\begin{aligned} \widehat{\mathbf{H}}_{\mathsf{ML}} &= \arg\min_{\mathbf{H}} \left\{ \|\mathbf{Y}_{\mathsf{P}} - \mathbf{H}\mathbf{X}_{\mathsf{P}}\|^2 \right\} \\ &= \mathbf{Y}_{\mathsf{P}}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}}(\mathbf{X}_{\mathsf{P}}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}})^{-1} \end{aligned} \tag{2.7}$$

or an MMSE criterion:

$$\begin{aligned} \widehat{\mathbf{H}}_{\mathsf{MMSE}} &= \mathbf{Y}_{\mathsf{P}} \left\{ \arg\min_{\mathbf{F}} \mathbb{E}\left[\|\mathbf{Y}_{\mathsf{P}}\mathbf{F} - \mathbf{H}\|^2 \right] \right\} \\ &= \mathbf{Y}_{\mathsf{P}}(\mathbf{X}_{\mathsf{P}}^{\mathsf{H}}\mathbf{M}\mathbf{X}_{\mathsf{P}} + rN_0\mathbf{I}_P)^{-1}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}}\mathbf{M} \end{aligned} \tag{2.8}$$

where it is defined

$$\mathbf{M} \triangleq \bar{\mathbf{H}}_0^{\mathsf{H}}\bar{\mathbf{H}}_0 + \frac{1}{K+1}\mathrm{Tr}\left(\mathbf{R}\right)\mathbf{T}.$$

under the separately-correlated fading assumption of eq. (2.2).

Remark 2.2.1. It is worth noting that the ML estimate (2.7) does not require the knowledge of the channel parameters $\bar{\mathbf{H}}_0$, $\mathbf{R}$ and $\mathbf{T}$ at the receiver whereas both the MMSE estimate (2.8) and the optimum receiver (as it will be shown in the following) need these parameters. Moreover, the existence of (2.7) requires the matrix $\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}$ to be invertible. Thus, the number of pilot intervals must satisfy the inequality $P \geq n_\mathsf{T}$, which is assumed to hold throughout the paper.

Remark 2.2.2. When the SNR grows very large (and hence $N_0 \to 0$), the ML and MMSE estimates are equivalent since

$$\mathbf{Y}_\mathsf{P}(\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M}\mathbf{X}_\mathsf{P})^{-1}\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M} = \mathbf{Y}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H})^{-1}$$

for every invertible matrix $\mathbf{M}$, provided that $\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}$ is invertible. In fact,

$$(\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M}\mathbf{X}_\mathsf{P})^{-1}\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M} = \mathbf{X}_\mathsf{P}^\mathsf{H}(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H})^{-1}$$

is equivalent to

$$\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M}(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}) = (\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{M}\mathbf{X}_\mathsf{P})\mathbf{X}_\mathsf{P}^\mathsf{H} \ ,$$

which is an identity from matrix product associativity.

After having obtained the (ML or MMSE) estimate $\widehat{\mathbf{H}}$, in the second stage the mismatched receiver uses it as if it were exactly equal to the channel matrix $\mathbf{H}$ and outputs the code word

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{MR}}(\mathbf{X})$$

where the corresponding *(ML or MMSE) mismatched metric* is defined as

$$\mu_{\mathsf{MR}}(\mathbf{X}) = \|\mathbf{Y} - \widehat{\mathbf{H}}\mathbf{X}\|^2 \ . \tag{2.9}$$

2.2.3 Optimum receiver

Contrary to the ML mismatched receiver, the optimum receiver does not require the explicit estimation of the channel matrix. Instead, it decodes the transmitted code word $\mathbf{X}$ by jointly processing the received signal matrices $\mathbf{Y}_\mathsf{P}$,$\mathbf{Y}$ and the pilot symbol matrix $\mathbf{X}_\mathsf{P}$. Assuming equally likely transmitted code words, the optimum receiver outputs the code word $\mathbf{X}$ maximizing

$$\begin{aligned}\widehat{\mathbf{X}} &= \arg\max_{\mathbf{X}} \Pr\left(\mathbf{X}|\mathbf{X}_\mathsf{P},\mathbf{Y},\mathbf{Y}_\mathsf{P}\right) \\ &= \arg\max_{\mathbf{X}} \frac{p\left(\mathbf{X},\mathbf{Y}\right) p\left(\mathbf{X}_\mathsf{P},\mathbf{Y}_\mathsf{P}\right)}{p\left(\mathbf{X}_\mathsf{P},\mathbf{Y},\mathbf{Y}_\mathsf{P}\right)} \\ &= \arg\max_{\mathbf{X}} \frac{\Pr\left(\mathbf{X}\right)\Pr\left(\mathbf{X}_\mathsf{P}\right)}{p\left(\mathbf{X}_\mathsf{P},\mathbf{Y},\mathbf{Y}_\mathsf{P}\right)} p\left(\mathbf{Y},\mathbf{Y}_\mathsf{P}|\mathbf{X},\mathbf{X}_\mathsf{P}\right) \\ &= \arg\max_{\mathbf{X}} \mathbb{E}_{\mathbf{H}}\left[p\left(\mathbf{Y},\mathbf{Y}_\mathsf{P}|\mathbf{X},\mathbf{X}_\mathsf{P},\mathbf{H}\right)\right] \\ &= \arg\max_{\mathbf{X}} \mathbb{E}_{\mathbf{H}}\left[p\left(\mathbf{Y}|\mathbf{X},\mathbf{H}\right) p\left(\mathbf{Y}_\mathsf{P}|\mathbf{X}_\mathsf{P},\mathbf{H}\right)\right] \ . \end{aligned} \tag{2.10}$$

Here, the conditional independence of $\mathbf{Y}$ and $\mathbf{Y}_\mathsf{P}$ given $\mathbf{X}$,$\mathbf{H}$ and $\mathbf{X}_\mathsf{P}$,$\mathbf{H}$, respectively, was used.

Then, the a posteriori probability in (2.10) can be calculated by writing explicitly the pdf's involved and then applying the result in Appendix A.2. As a result, the following relation can be obtained

$$\begin{aligned}p\left(\mathbf{Y},\mathbf{Y}_\mathsf{P}|\mathbf{X},\mathbf{X}_\mathsf{P}\right) =& (\pi N_0)^{-(N+P)n_\mathsf{R}} \exp\left(-\frac{\|\mathbf{Y}_\mathsf{P}-\bar{\mathbf{H}}_0\mathbf{X}_\mathsf{P}\|^2+\|\mathbf{Y}-\bar{\mathbf{H}}_0\mathbf{X}\|^2}{N_0}\right) \\ &\cdot \mathbb{E}_{\mathbf{W}}\Big\{\mathrm{etr}\left(-(\mathbf{W}\mathbf{A}(\mathbf{X})\mathbf{W}^\mathsf{H}\mathbf{R}+\mathbf{W}\mathbf{B}(\mathbf{X})^\mathsf{H}+\mathbf{B}(\mathbf{X})\mathbf{W}^\mathsf{H})\right)\Big\}\end{aligned} \tag{2.11}$$

where

$$\mathbf{A}(\mathbf{X}) \triangleq \frac{\mathbf{T}^{1/2}(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}+\mathbf{X}\mathbf{X}^\mathsf{H})\mathbf{T}^{1/2}}{N_0(K+1)} \tag{2.12}$$

$$\mathbf{B}(\mathbf{X}) \triangleq \frac{\mathbf{R}^{1/2}[(\mathbf{Y}_\mathsf{P}-\bar{\mathbf{H}}_0\mathbf{X}_\mathsf{P})\mathbf{X}_\mathsf{P}^\mathsf{H}+(\mathbf{Y}-\bar{\mathbf{H}}_0\mathbf{X})\mathbf{X}^\mathsf{H}]\mathbf{T}^{1/2}}{N_0\sqrt{K+1}} \tag{2.13}$$

Finally, from Theorem A.2.1 of Appendix A.2 the a posteriori probability can be derived:

$$\begin{aligned}p\left(\mathbf{Y},\mathbf{Y}_\mathsf{P}|\mathbf{X},\mathbf{X}_\mathsf{P}\right) =& (\pi N_0)^{-(N+P)n_\mathsf{R}} \det(\mathbf{I}_{n_\mathsf{T}n_\mathsf{R}}+\mathbf{C}(\mathbf{X}))^{-1} \\ &\cdot \exp\left(-\frac{\|\mathbf{Y}_\mathsf{P}-\bar{\mathbf{H}}_0\mathbf{X}_\mathsf{P}\|^2+\|\mathbf{Y}-\bar{\mathbf{H}}_0\mathbf{X}\|^2}{N_0}\right) \\ &\cdot \exp\left(\mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}^\mathsf{H}(\mathbf{I}_{n_\mathsf{T}n_\mathsf{R}}+\mathbf{C}(\mathbf{X}))^{-1}\mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}\right)\end{aligned} \tag{2.14}$$

where $\mathbf{C}(\mathbf{X}) \triangleq \mathbf{A}(\mathbf{X})^{\mathsf{T}} \otimes \mathbf{R}$. Hence, the decision metric for the optimum receiver is obtained as the negative logarithm of (2.14) which yields, after proper scaling and dropping terms independent of $\mathbf{X}$,

$$\begin{aligned}\mu_{\mathsf{OR}}(\mathbf{X}) =& \|\mathbf{Y} - \bar{\mathbf{H}}_0\mathbf{X}\|^2 + N_0 \ln\det(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}} + \mathbf{C}(\mathbf{X})) \\ &- N_0 \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}^{\mathsf{H}} (\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}} + \mathbf{C}(\mathbf{X}))^{-1} \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\} \end{aligned} \tag{2.15}$$

so that the optimum receiver outputs the code word

$$\widehat{\mathbf{X}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{OR}}(\mathbf{X}) \ .$$

It can be seen that, after proper interpretation of the system parameters, the metric (2.15) specializes to [36, eq. (20)] in the case of independent Rayleigh fading and to the logarithm of [42, eq. (4)] in the absence of pilot symbols.

2.2.4 Iterative metric computation

The computational complexity of exhaustive decoding based on the metric (2.15) grows exponentially with the code word length N. However, it can be considerably reduced by resorting to an iterative algorithm.

Following the approach suggested in [36], the metric (2.15) is split into two parts, the former representing the current decoding state and the latter accounting for the state transition. This approach applies directly to trellis STCs [5].

The following decompositions are defined:

$$\mathbf{X} = (\mathbf{X}^-, \mathbf{x}) \qquad \text{and} \qquad \mathbf{Y} = (\mathbf{Y}^-, \mathbf{y}) \ ,$$

where $\mathbf{X}^-$ denotes the part of the code word ranging over all the preceding symbol times, $\mathbf{x}$ is the current symbol vector, $\mathbf{Y}^-$ denotes the part of the received word ranging over all the preceding symbol times, and $\mathbf{y}$ is the current received vector.

Then, the optimum decision metric can be written as

$$\mu_{\mathsf{OR}}(\mathbf{X}) = \mu_{\mathsf{OR}}(\mathbf{X}^-) + \Delta\mu_{\mathsf{OR}}(\mathbf{X}^-,\mathbf{x})$$

where the metric increment (corresponding to a branch metric in the Viterbi decoding algorithm) is given by (see Appendix A.3):

$$\begin{aligned}\Delta\mu_{\mathsf{OR}}(\mathbf{X}^-,\mathbf{x}) \triangleq& \|\mathbf{y} - \bar{\mathbf{H}}_0\mathbf{x}\|^2 \\ &+ N_0 \ln\det\{\mathbf{I}_{n_{\mathsf{R}}} + \mathbf{\Psi}(\mathbf{x})^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X}^-)\mathbf{\Psi}(\mathbf{x})\} \\ &- N_0 \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}^{\mathsf{H}} \mathbf{\Lambda}(\mathbf{X}) \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\} \\ &+ N_0 \mathrm{vec}\left\{\mathbf{B}(\mathbf{X}^-)\right\}^{\mathsf{H}} \mathbf{\Lambda}(\mathbf{X}^-) \mathrm{vec}\left\{\mathbf{B}(\mathbf{X}^-)\right\} \end{aligned} \tag{2.16}$$

where (subscripts refer to matrix sizes)

$$\begin{cases} \boldsymbol{\Psi}(\mathbf{x})_{n_\mathsf{T} n_\mathsf{R} \times n_\mathsf{R}} & \triangleq \dfrac{(\mathbf{T}^{1/2}\mathbf{x})^*}{\sqrt{N_0(K+1)}} \otimes \mathbf{R}^{1/2} \\ \boldsymbol{\Lambda}(\mathbf{X})_{n_\mathsf{T} n_\mathsf{R} \times n_\mathsf{T} n_\mathsf{R}} & \triangleq \boldsymbol{\Lambda}(\mathbf{X}^-) + \Delta\boldsymbol{\Lambda}(\mathbf{X}^-,\mathbf{x}) \\ \Delta\boldsymbol{\Lambda}(\mathbf{X}^-,\mathbf{x})_{n_\mathsf{T} n_\mathsf{R} \times n_\mathsf{T} n_\mathsf{R}} & \triangleq -\boldsymbol{\Lambda}(\mathbf{X}^-)\boldsymbol{\Psi}(\mathbf{x})[\mathbf{I}_{n_\mathsf{R}} + \boldsymbol{\Psi}(\mathbf{x})^\mathsf{H}\boldsymbol{\Lambda}(\mathbf{X}^-)\boldsymbol{\Psi}(\mathbf{x})]^{-1}\boldsymbol{\Psi}(\mathbf{x})^\mathsf{H}\boldsymbol{\Lambda}(\mathbf{X}^-) \\ \mathbf{B}(\mathbf{X})_{n_\mathsf{R} \times n_\mathsf{T}} & \triangleq \mathbf{B}(\mathbf{X}^-) + \Delta\mathbf{B}(\mathbf{x}) \\ \Delta\mathbf{B}(\mathbf{x})_{n_\mathsf{R} \times n_\mathsf{T}} & \triangleq \mathbf{R}^{1/2}\dfrac{\mathbf{y} - \bar{\mathbf{H}}_0\mathbf{x}}{N_0\sqrt{K+1}}\mathbf{x}^\mathsf{H}\mathbf{T}^{1/2} \end{cases}$$

At the beginning, the decoding algorithm initializes the matrices $\boldsymbol{\Lambda}(\mathbf{X})$ and $\mathbf{B}(\mathbf{X})$ assuming that $\mathbf{X}$ is the empty matrix $\varnothing$. Thus, the following initialization rules hold:

$$\boldsymbol{\Lambda}(\varnothing) = \left(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \frac{(\mathbf{T}^{1/2}\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{T}^{1/2})^\mathsf{T}}{N_0(K+1)} \otimes \mathbf{R}\right)^{-1}$$

$$\mathbf{B}(\varnothing) = \frac{\mathbf{R}^{1/2}(\mathbf{Y}_\mathsf{P} - \bar{\mathbf{H}}_0\mathbf{X}_\mathsf{P})\mathbf{X}_\mathsf{P}^\mathsf{H}\mathbf{T}^{1/2}}{N_0\sqrt{K+1}}$$

2.2.5 Complexity

A fair system comparison requires to consider the complexity involved in both receiver implementations. With trellis space–time coding, both receivers are based on the Viterbi decoding algorithm and hence complexity depends essentially on the branch metric computation.

Here the complexities of the ML mismatched and optimum receivers are compared. We recall here some fundamental properties:

1. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, both sized $a \times b$, their sum $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ requires ab complex sums.

2. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, sized $a \times b$ and $b \times c$, respectively, their product $\mathbf{Z} = \mathbf{XY}$ requires acb complex products and $ac(b-1)$ complex sums.

3. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, sized $a \times b$ and $c \times d$, respectively, their Kronecker product $\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y}$ requires $abcd$ complex products.

4. The solution of a set of a linear equations in a unknowns, i.e. the solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ of $\mathbf{b} = \mathbf{Ax}$, where $\mathbf{x}$ contains the a unknowns, $\mathbf{A}$, sized $a \times a$, is the matrix of the coefficients and the vector $\mathbf{b}$ contains the a right–hand side quantities, performed by LU decomposition of $\mathbf{A}$ followed by forward and backward substitutions, requires $\frac{a^3}{3}$ complex products and $\frac{a^3}{3}$ complex sums for the LU decomposition, while the forward and backward substitutions require a^2 complex products and a^2 complex sums [57].

5. Hence, the solution $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ of $\mathbf{B} = \mathbf{A}\mathbf{X}$, with $\mathbf{X}$ and $\mathbf{B}$ sized $a \times b$, requires $\frac{a^3}{3} + ba^2$ complex products and $\frac{a^3}{3} + ba^2$ complex sums.

6. Moreover, once obtained the LU decomposition of $\mathbf{A}$, the evaluation of the determinant $|\mathbf{A}|$ requires $a - 1$ complex products.

7. Given a vector $\mathbf{x}$ of a elements its norm $\|\mathbf{x}\|^2$ requires a complex products and $a - 1$ complex sums.

8. A complex sum requires 2 real sums.

9. A complex product requires 4 real products and 2 real sums.

Under the quasi-static fading assumption, the ML mismatched receiver requires only the estimation of the channel matrix whereas the optimum receiver requires also the estimation of the statistic parameters $\mathbf{H}_0$, K, $\mathbf{R}$, and $\mathbf{T}$. Therefore, for the mismatched receiver the channel matrix is estimated at the beginning of the frame decoding by (2.7), whose computational cost amounts essentially to a matrix multiplication since $\mathbf{X}_\mathsf{P}^\mathsf{H}(\mathbf{X}_\mathsf{P}\mathbf{X}_\mathsf{P}^\mathsf{H})^{-1}$ can be calculated off-line.

The dominant complexity is the branch complexity which amounts to $n_\mathsf{T} n_\mathsf{R}$ complex products and $n_\mathsf{T} n_\mathsf{R}$ complex sums per branch metric as n_T,$n_\mathsf{R} \to \infty$. The details can be found in Table 2.1.

Table 2.1. Computational complexity of mismatched receiver branch metric evaluation (2.9)

No.	**Operation**	**Complex products No.**	**Complex sums No.**
(1)	$\widehat{\mathbf{H}}\mathbf{x}$	$n_\mathsf{T} n_\mathsf{R}$	$(n_\mathsf{T} - 1)n_\mathsf{R}$
(2)	$\mathbf{y} - (1)$	-	n_R
(3)	$\|(2)\|^2$	n_R	$n_\mathsf{R} - 1$
Dominant complexity		$n_\mathsf{T} n_\mathsf{R}$	$n_\mathsf{T} n_\mathsf{R}$

The implementation of the optimum receiver entails a higher complexity because of the computation of metric (2.16). Table 2.2 summarizes the complexities of the involved operations.

Thus, as n_T,$n_\mathsf{R} \to \infty$, the overall complexity amounts to $2n_\mathsf{T}^2 n_\mathsf{R}^3$ operations per branch metric so that the increase in complexity required by the optimum receiver is $2n_\mathsf{T} n_\mathsf{R}^2$.

Table 2.2. Computational complexity of optimum receiver branch metric evaluation (2.16)

No.	Operation	Complex products No.	Complex sums No.
(1)	$\|\mathbf{y}-\tilde{\mathbf{H}}_0\mathbf{x}\|^2$	$(n_\mathsf{T}+1)n_\mathsf{R}$	$(n_\mathsf{T}+1)n_\mathsf{R}-1$
	$\boldsymbol{\Psi}(\mathbf{x})$		
(2)	$\mathbf{T}^{1/2}\mathbf{x}$	n_T^2	$n_\mathsf{T}(n_\mathsf{T}-1)$
(3)	$(2)\otimes\mathbf{R}$	$n_\mathsf{T}n_\mathsf{R}^2$	–
	$\Delta\boldsymbol{\Lambda}(\mathbf{X}^-,\mathbf{x})$		
(4)	$\boldsymbol{\Lambda}(\mathbf{X}^-)\boldsymbol{\Psi}(\mathbf{x})$	$n_\mathsf{T}^2n_\mathsf{R}^3$	$n_\mathsf{T}(n_\mathsf{T}n_\mathsf{R}-1)n_\mathsf{R}^2$
(5)	$\mathbf{I}_{n_\mathsf{R}}+\boldsymbol{\Psi}(\mathbf{x})^\mathsf{H}(4)$	$n_\mathsf{T}n_\mathsf{R}^3$	$n_\mathsf{T}n_\mathsf{R}^3$
(6)	$(5)^{-1}(4)^\mathsf{H}$	$\frac{n_\mathsf{R}^3}{3}+n_\mathsf{T}n_\mathsf{R}^3$	$\frac{n_\mathsf{R}^3}{3}+n_\mathsf{T}n_\mathsf{R}^3$
(7)	(4)(6)	$n_\mathsf{T}^2n_\mathsf{R}^3$	$n_\mathsf{T}^2(n_\mathsf{R}-1)n_\mathsf{R}^2$
(8)	$\boldsymbol{\Lambda}(\mathbf{X})=\boldsymbol{\Lambda}(\mathbf{X}^-)+(7)$	–	$n_\mathsf{T}^2n_\mathsf{R}^2$
	$\Delta\mathbf{B}(\mathbf{x})$		
(9)	$\mathbf{R}^{1/2}(1)$	n_R^2	$n_\mathsf{R}(n_\mathsf{R}-1)$
(10)	$(9)(2)^\mathsf{H}$	$n_\mathsf{T}n_\mathsf{R}$	–
(11)	$\mathbf{B}(\mathbf{X})=\mathbf{B}(\mathbf{X}^-)+(10)$	–	$n_\mathsf{T}n_\mathsf{R}$
(12)	$\det\{(5)\}$	$n_\mathsf{R}-1$	–
(13)	$\mathrm{vec}\{(11)\}^\mathsf{H}(8)$	$n_\mathsf{T}^2n_\mathsf{R}^2$	$n_\mathsf{T}n_\mathsf{R}(n_\mathsf{T}n_\mathsf{R}-1)$
(14)	$(13)\mathrm{vec}\{(11)\}$	$n_\mathsf{T}n_\mathsf{R}$	$n_\mathsf{T}n_\mathsf{R}-1$
Dominant complexity		$2n_\mathsf{T}^2n_\mathsf{R}^3$	$2n_\mathsf{T}^2n_\mathsf{R}^3$

2.2.6 Parameter estimation

The optimum receiver described in Section 2.2.3 requires to know several parameters of the statistic distribution of the channel matrix $\mathbf{H}$. More precisely, the receiver needs the average channel matrix $\bar{\mathbf{H}}_0$, the Rice factor K, and the correlation matrices $\mathbf{R}$ and $\mathbf{T}$.

In the earlier sections these parameters were assumed to be known exactly but a practical receiver has to estimate them and one way to obtain the estimate is to use the pilot symbols transmitted within the data frame. More precisely, it is proposed to estimate the channel parameters by resorting to the ML estimates of the channel matrix $\mathbf{H}$ obtained over different fading blocks.

Let us write the channel equation corresponding to the ith fading block as

$$\mathbf{Y}_{\mathsf{P}}^{(i)} = \mathbf{H}^{(i)}\mathbf{X}_{\mathsf{P}} + \mathbf{Z}_{\mathsf{P}}^{(i)} \qquad\qquad i = 1,2,\ldots$$

As already seen in Section 2.2.2, the corresponding ML estimate of $\mathbf{H}^{(i)}$ is given by

$$\widehat{\mathbf{H}}^{(i)} = \mathbf{Y}_{\mathsf{P}}^{(i)}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}}(\mathbf{X}_{\mathsf{P}}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}})^{-1} = \mathbf{H}^{(i)} + \mathbf{Z}_{\mathsf{P}}^{(i)}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}}(\mathbf{X}_{\mathsf{P}}\mathbf{X}_{\mathsf{P}}^{\mathsf{H}})^{-1} \ .$$

Since the joint ML estimation of all these parameters is quite complex, it is preferable to resort to the following simple suboptimal approach.

Assuming that L channel matrix estimates

$$\widehat{\mathbf{H}}^{(1)},\ldots,\widehat{\mathbf{H}}^{(L)}$$

are available, it is possible to estimate the statistic parameters as follows:

- The channel mean is ML estimated by

$$\widehat{\mathbf{H}}_0 = \frac{1}{L}\sum_{i=1}^{L}\mathbf{H}^{(i)} \ .$$

- The Rice factor K is estimated by

$$\widehat{K} = \frac{\|\widehat{\mathbf{H}}_0\|^2}{\dfrac{1}{L}\displaystyle\sum_{i=1}^{L}\|\mathbf{H}^{(i)} - \widehat{\mathbf{H}}_0\|^2} \tag{2.17}$$

- Next, $\mathbf{T}$ and $\mathbf{R}$ are estimated by noting that

$$\mathbb{E}\left[\mathbf{H}\mathbf{H}^{\mathsf{H}}\right] = \bar{\mathbf{H}}_0\bar{\mathbf{H}}_0^{\mathsf{H}} + \frac{1}{K+1}\mathrm{Tr}\left(\mathbf{T}\right)\mathbf{R}$$
$$\mathbb{E}\left[\mathbf{H}^{\mathsf{H}}\mathbf{H}\right] = \bar{\mathbf{H}}_0^{\mathsf{H}}\bar{\mathbf{H}}_0 + \frac{1}{K+1}\mathrm{Tr}\left(\mathbf{R}\right)\mathbf{T}$$
$$\mathbb{E}\left[\|\mathbf{H}\|^2\right] = \|\bar{\mathbf{H}}_0\|^2 + \frac{1}{K+1}\mathrm{Tr}\left(\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{T}\right)$$

Hence, set

$$\hat{q} = \frac{\widehat{K}+1}{L}\sum_{i=1}^{L}\left(\|\widehat{\mathbf{H}}^{(i)}\|^2 - \|\widehat{\mathbf{H}}_0\|^2\right)$$
$$\widehat{\mathbf{R}} = \frac{\widehat{K}+1}{\sqrt{\hat{q}}\;L}\sum_{i=1}^{L}\left(\widehat{\mathbf{H}}^{(i)}\left(\widehat{\mathbf{H}}^{(i)}\right)^{\mathsf{H}} - \widehat{\mathbf{H}}_0\widehat{\mathbf{H}}_0^{\mathsf{H}}\right)$$
$$\widehat{\mathbf{T}} = \frac{\widehat{K}+1}{\sqrt{\hat{q}}\;L}\sum_{i=1}^{L}\left(\left(\widehat{\mathbf{H}}^{(i)}\right)^{\mathsf{H}}\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0^{\mathsf{H}}\widehat{\mathbf{H}}_0\right)$$

after changing statistical with arithmetic means and dividing the estimate $\hat{q}$ of $\mathrm{Tr}\left(\mathbf{T}\right)\mathrm{Tr}\left(\mathbf{R}\right)$ evenly between $\mathbf{T}$ and $\mathbf{R}$. Notice that the two matrices $\widehat{\mathbf{T}}$ and $\widehat{\mathbf{R}}$ are positive definite since

$$\frac{1}{L}\sum_{i=1}^{L}\left(\widehat{\mathbf{H}}^{(i)}\left(\widehat{\mathbf{H}}^{(i)}\right)^{\mathsf{H}} - \widehat{\mathbf{H}}_0\widehat{\mathbf{H}}_0^{\mathsf{H}}\right) = \frac{1}{L}\sum_{i=1}^{L}(\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0)\left(\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0\right)^{\mathsf{H}}$$

and

$$\frac{1}{L}\sum_{i=1}^{L}\left(\left(\widehat{\mathbf{H}}^{(i)}\right)^{\mathsf{H}}\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0^{\mathsf{H}}\widehat{\mathbf{H}}_0\right) = \frac{1}{L}\sum_{i=1}^{L}\left(\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0\right)^{\mathsf{H}}(\widehat{\mathbf{H}}^{(i)} - \widehat{\mathbf{H}}_0)\ ,$$

and by the definition $\widehat{\mathbf{H}}_0 = \frac{1}{L}\sum_{i=1}^{L}\widehat{\mathbf{H}}^{(i)}$.

The receiver will then operate according to two different modes:

Training mode. The first L blocks are used to estimate the channel parameters in parallel to the normal operation of the mismatched receiver.

Operating mode. The next blocks are processed by using the optimum receiver with the parameters estimated during all the fading blocks preceding the current one.

Remark 2.2.3. Our definition of training and operating modes applies to full channel frames and not to the parts of the frame assigned to pilots and to data, respectively.

It must be stressed the fact that user data are transmitted *during both modes*, so that there is no loss in data rate with respect to the ideal case. However, the FER is expected to be worse during the training mode (when the mismatched receiver is used) than during the operating mode (when the optimum receiver based on the estimated channel parameters is used).

Since parameter estimation is performed not only during the training mode but also during the operating mode, the receiver becomes insensitive to the duration of the training mode. In fact, if the training mode was too short, bad parameter estimation would affect only the first few frames of the operating mode, until parameter estimation improved enough to reach a steady-state performance. Simulation results show that parameter estimation converges rather quickly in the cases considered.

Remark 2.2.4. Notice that the number of parameters to estimate (in $\mathbf{R}$ and $\mathbf{T}$) grows as $n_{\mathsf{R}}^2 + n_{\mathsf{T}}^2$, which may be large for large number of antennas (where using the prior knowledge on antenna correlation makes a larger difference, presumably). As a result, the duration of the training mode has to grow roughly proportionally with $\max\{n_{\mathsf{T}},n_{\mathsf{R}}\}$. In our numerical examples, this effect cannot be noticed since a 2×2 MIMO system (i.e., a system with $n_{\mathsf{T}} = 2$ transmit and $n_{\mathsf{R}} = 2$ receive antennas) is considered.

2.3 Numerical Results

In this section a MIMO system with $n_{\mathsf{T}} = 2$ transmit and $n_{\mathsf{R}} = 2$ receive antennas is considered.

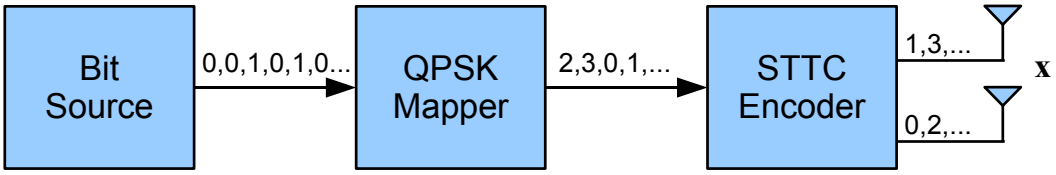


Figure 2.2. The narrowband transmitting scheme

The space-time code used for channel coding is obtained from an optimum rate-2/4 binary convolutional code with generator matrix [58]

$$\mathbf{G}_{STC} = \begin{pmatrix} 0 & 3 & 1 & 2 \\ 3 & 1 & 2 & 1 \end{pmatrix}$$

The resulting code trellis is illustrated in Figure 2.3.

Every vector of four encoded bits is grouped in two quaternary symbols and mapped to QPSK symbols by the rule $x \mapsto \sqrt{E_{\mathsf{S}}} \exp(jx\pi/2)$, where E_{S} is the average symbol energy (see Figure 2.4).

The complex symbols are then transmitted by the two transmit antennas.

The frame length, including trellis termination, is $N = 130$. The pilot symbol energy is assumed to be equal to the data symbol energy, i.e., $E_{\mathsf{P}} = E_{\mathsf{S}}$. In order to compare variable pilot symbol settings, the average symbol energy E_{S} is obtained from the information bit energy E_b by taking into account the loss in efficiency due to pilot symbol insertion [36]:

$$E_{\mathsf{S}} = \frac{N}{N+P} \frac{\mu_b E_b}{n_{\mathsf{T}}}$$

where μ_b denotes the number of information bits per symbol transmitted by the space-time code (in this case, $\mu_b = 2$). It is also assumed that the pilot symbol matrix $\mathbf{X}_{\mathsf{P}}$ is orthogonal, namely,

$$\mathbf{X}_{\mathsf{P}} \mathbf{X}_{\mathsf{P}}^{\mathsf{H}} = P E_{\mathsf{P}} \mathbf{I}_{n_{\mathsf{T}}}$$

Simulations have been carried out for the mismatched and optimum receivers under different fading assumptions:

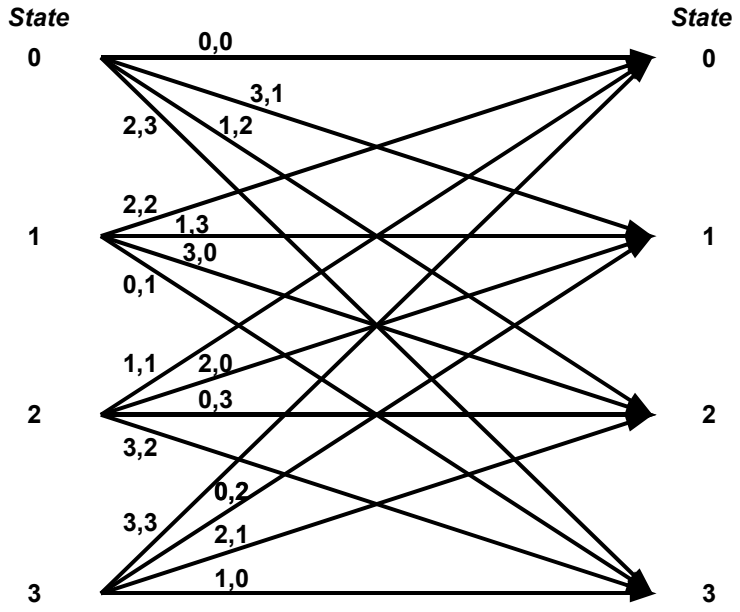


Figure 2.3. Branches labels are given by symbol pairs, the former transmitted from antenna 1, the latter from antenna 2

- zero-mean (Rayleigh) correlated,
- nonzero mean (Rice) uncorrelated,
- nonzero mean (Rice) correlated.

Correlation is assumed to be separate and modeled by the transmit and receive correlation matrices $\mathbf{T}$ and $\mathbf{R}$, which are given according to the exponential model proposed in [11, 12]. Then,

$$(\mathbf{T})_{ij} = \begin{cases} \rho_{\mathsf{T}}^{j-i}, & i \leq j \\ (\mathbf{T})_{ji}^{*}, & i > j \end{cases}$$

and

$$(\mathbf{R})_{ij} = \begin{cases} \rho_{\mathsf{R}}^{j-i}, & i \leq j \\ (\mathbf{R})_{ji}^{*}, & i > j \end{cases}$$

where ρ_{T} and ρ_{R} are the complex correlation coefficients of neighboring transmit and receive antennas, respectively.

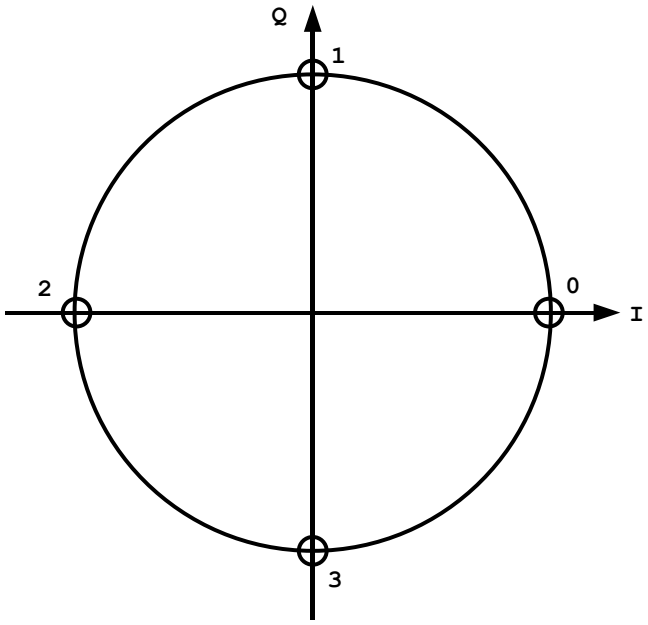


Figure 2.4. QPSK mapping

Following [59, 17, 13], the nonzero channel matrix mean is assumed to have unit rank and is given by

$$\mathbf{H}_0 = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}_{n_\mathsf{R} \times n_\mathsf{T}} .$$

Performance results are reported in terms of FER versus E_b/N_0 in dB.

In the following a MIMO system with n_T transmit and n_R receive antenna is referred to as an $n_\mathsf{T} \times n_\mathsf{R}$ MIMO system.

2.3.1 Optimization of the number of pilot symbol intervals per frame

Our first set of results considers the effect of the number of pilot symbol intervals per frame, P, on the FER performance under different operating conditions and fixed E_b/N_0.

Specifically, it is assumed that $E_b/N_0 = 10$ dB and consider the cases of $K = 0$ (Rayleigh, Figure 2.5), 0 dB (Rice, Figure 2.6), 10 dB (RiceFigure 2.7) and 20 dB (RiceFigure 2.8).

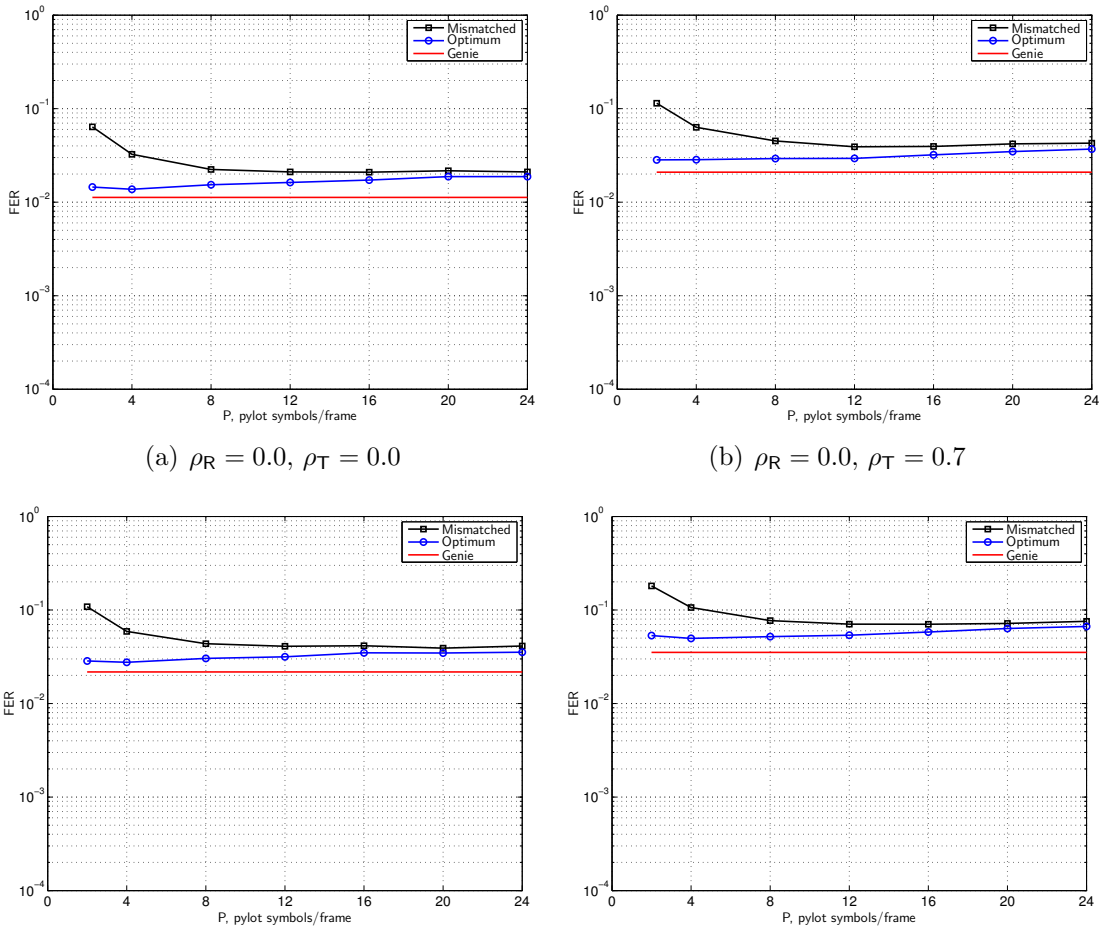
In all cases four correlation settings are considered:

1. $\rho_{\mathsf{T}} = 0$,$\rho_{\mathsf{R}} = 0$ (uncorrelated case);
2. $\rho_{\mathsf{T}} = 0.7$,$\rho_{\mathsf{R}} = 0$ (transmit-only correlation);
3. $\rho_{\mathsf{T}} = 0$,$\rho_{\mathsf{R}} = 0.7$ (receive-only correlation);
4. $\rho_{\mathsf{T}} = 0.7$,$\rho_{\mathsf{R}} = 0.7$ (joint transmit and receive correlation).

Analyzing the results reported in Figs 2.5, 2.6, 2.7 and 2.8 it can be noticed that the optimum value of P for the mismatched receiver depends somewhat on the Rice factor and correlation conditions but in most cases is close to 16.

Similarly, the best P for the optimum receiver is equal to 4 (Rayleigh case and Rice case for $K = 0$ dB and no receive correlation and $K = 0$ dB with $\rho_{\mathsf{T}} = 0.0$) or 2 (in the other cases, especially Rice case with $K = 20$ dB).

The diagrams confirm that the optimum receiver outperforms the mismatched one in all cases and its FER is lower bounded by that of the *genie* receiver, i.e., a receiver with perfect CSI provided by some genie device exactly (and not by pilot-aided estimation).



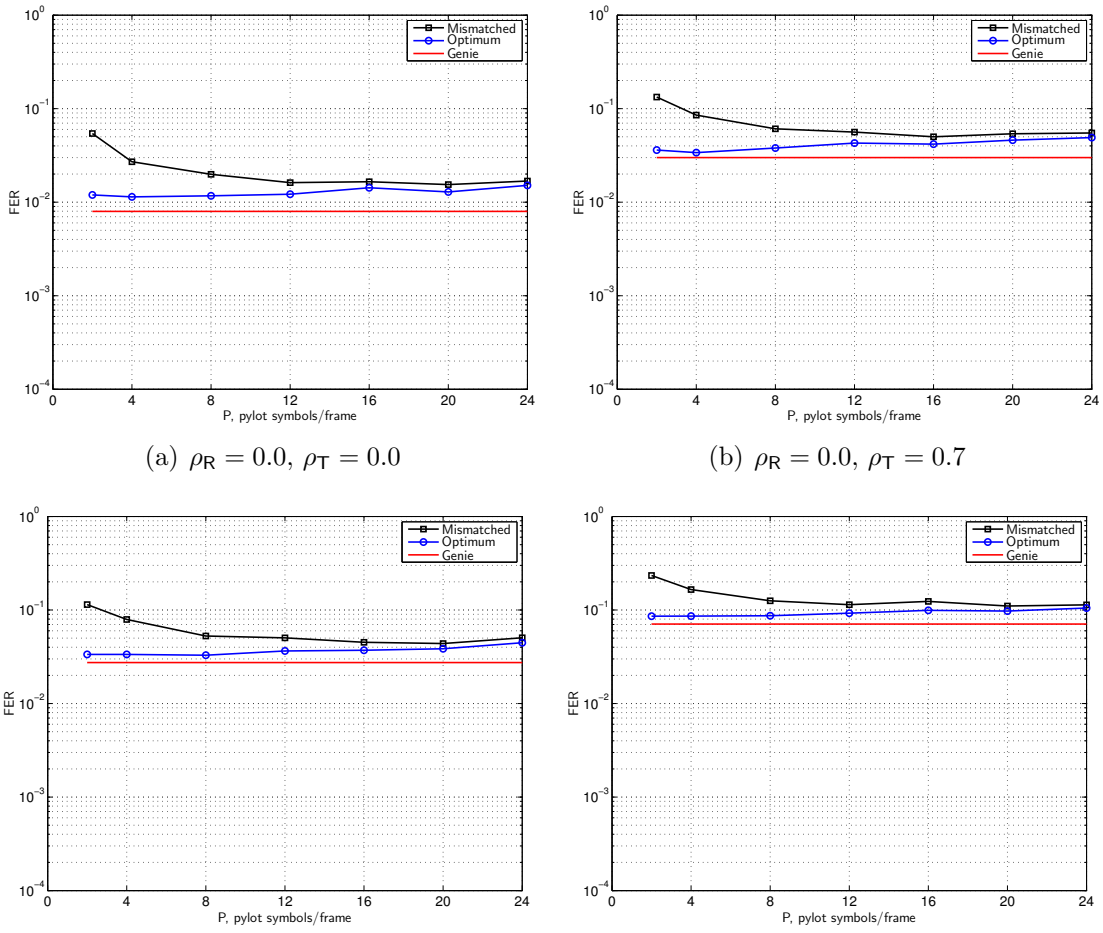
(a) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.0$

(b) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.7$

(c) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.0$

(d) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.7$

Figure 2.5. Plot of FER versus number of pilot symbol intervals per frame for a 2×2 MIMO system with $E_b/N_0 = 10$ dB and $K = 0$ (Rayleigh)



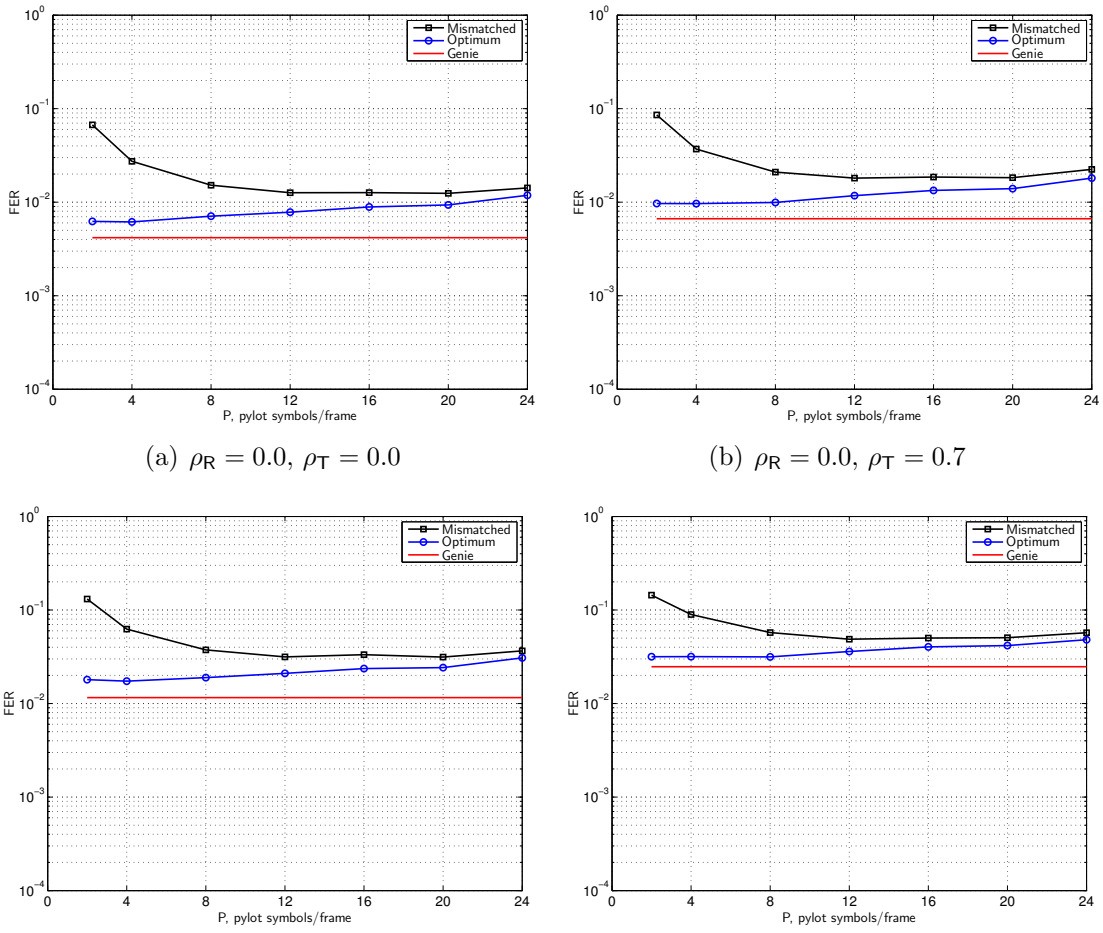
(a) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.0$

(b) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.7$

(c) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.0$

(d) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.7$

Figure 2.6. Plot of FER versus number of pilot symbol intervals per frame for a 2×2 MIMO system with $E_b/N_0 = 10$ dB and $K = 0$ dB (Rice)



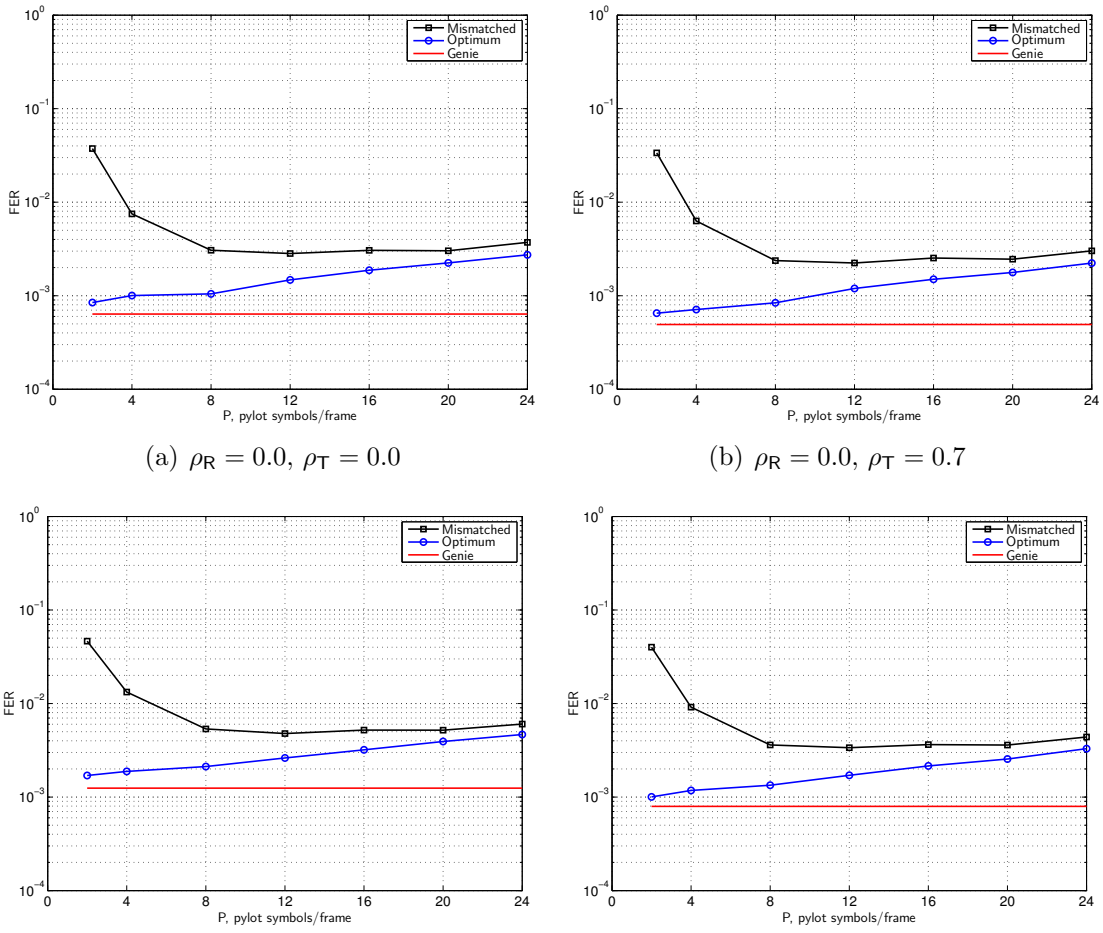
(a) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.0$

(b) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.7$

(c) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.0$

(d) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.7$

Figure 2.7. Plot of FER versus number of pilot symbol intervals per frame for a 2×2 MIMO system with $E_b/N_0 = 10$ dB and $K = 10$ dB (Rice)



(a) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.0$

(b) $\rho_{\mathsf{R}} = 0.0$, $\rho_{\mathsf{T}} = 0.7$

(c) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.0$

(d) $\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.7$

Figure 2.8. Plot of FER versus number of pilot symbol intervals per frame for a 2×2 MIMO system with $E_b/N_0 = 10$ dB and $K = 20$ dB (Rice)

2.3.2 Optimum receiver with Rayleigh fading: Effects of correlation

The effects of correlation are investigated by considering the case of zero-mean channel matrix ($K = 0$), i.e., the correlated Rayleigh fading case. The correlation settings specified in the previous section are considered setting P equal to its optimum value ($P = 4$ for the 2×2 system).

Figure 2.9 shows the performance of a 2×2 MIMO system with the optimum receiver and $P = 4$ (optimum value). It can be noticed that, with either receive-only or transmit-only correlation, the performance degradation is about 1.2 dB. With joint transmit and receive correlation, there is an additional degradation of about 1 dB. The curves also show that, at high values of E_b/N_0, the transmit-only correlated system FER is better than in the receive-only correlated case, in accordance with the observations reported in [56, Simulation Example 2].

Figure 2.10 illustrates the joint effect of the two correlation coefficients for a 2×2 MIMO system with $K = 0$, $P = 4$, and $E_b/N_0 = 10$ dB by reporting constant-FER level curves. The curves are almost symmetric with respect to the quadrant bisector, implying that both transmit and receive correlation coefficients have a similar impact on the system performance. However, it is expected that this symmetry disappears when the E_b/N_0 ratio becomes larger, in accordance with [56, Simulation Example 2]. Moreover, it can be noticed that the performance degradation due to correlation increases steeply when the correlation coefficients exceed 0.5.

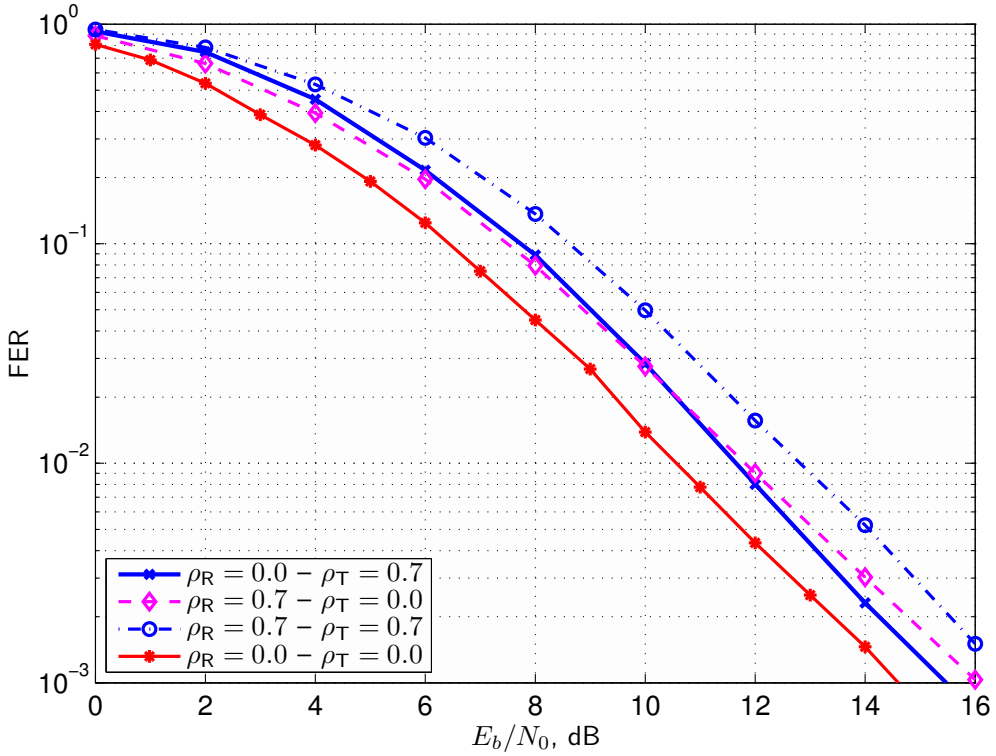


Figure 2.9. FER of a 2×2 Rayleigh fading MIMO system versus E_b/N_0. Optimum receiver with $P = 4$

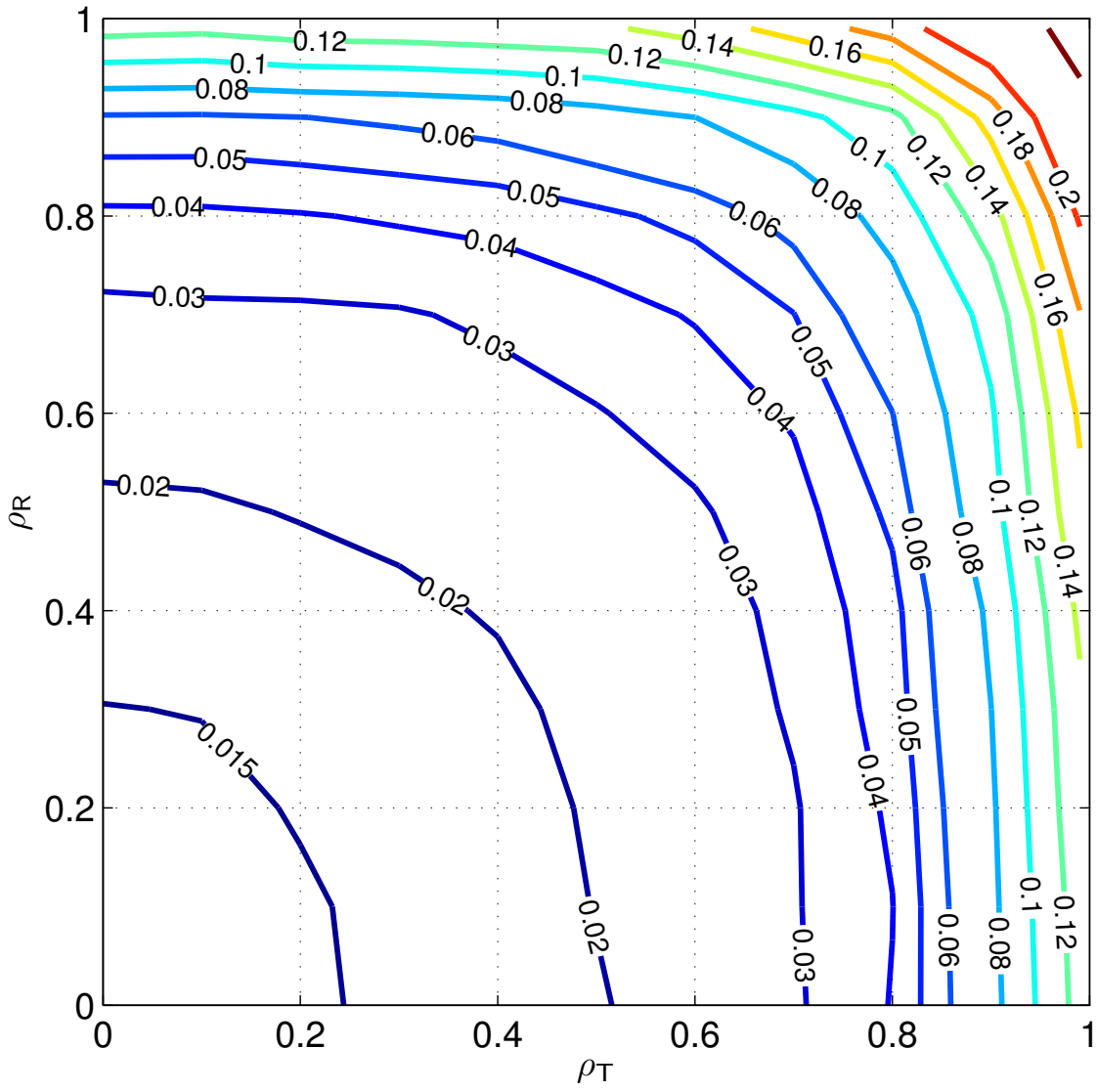


Figure 2.10. FER of a 2×2 Rayleigh fading MIMO system versus ρ_R, ρ_T. $E_b/N_0 = 10$ dB. Optimum receiver with $P = 4$

2.3.3 Optimum receiver with Rice fading: Effects of correlation and the Rice factor

Focusing on a 2×2 MIMO system the effect of the Rice factor on the system performance is analyzed in the uncorrelated case and in the case of joint transmit and receive correlation ($\rho_{\mathsf{R}} = 0.7$, $\rho_{\mathsf{T}} = 0.7$). Figure 2.11 shows the FER performance for $K = 0$ dB and 20 dB. It can be noticed that correlation has a practically no effect on the FER at high K since the channel is almost unfaded. On the contrary, correlation has a considerable impact at $K = 0$ dB where the correlation penalty amounts to approximately 3.5 dB.

Moreover, these results show that the FER performance improves by increasing the Rice factor K as far as the E_b/N_0 ratio is sufficiently high.

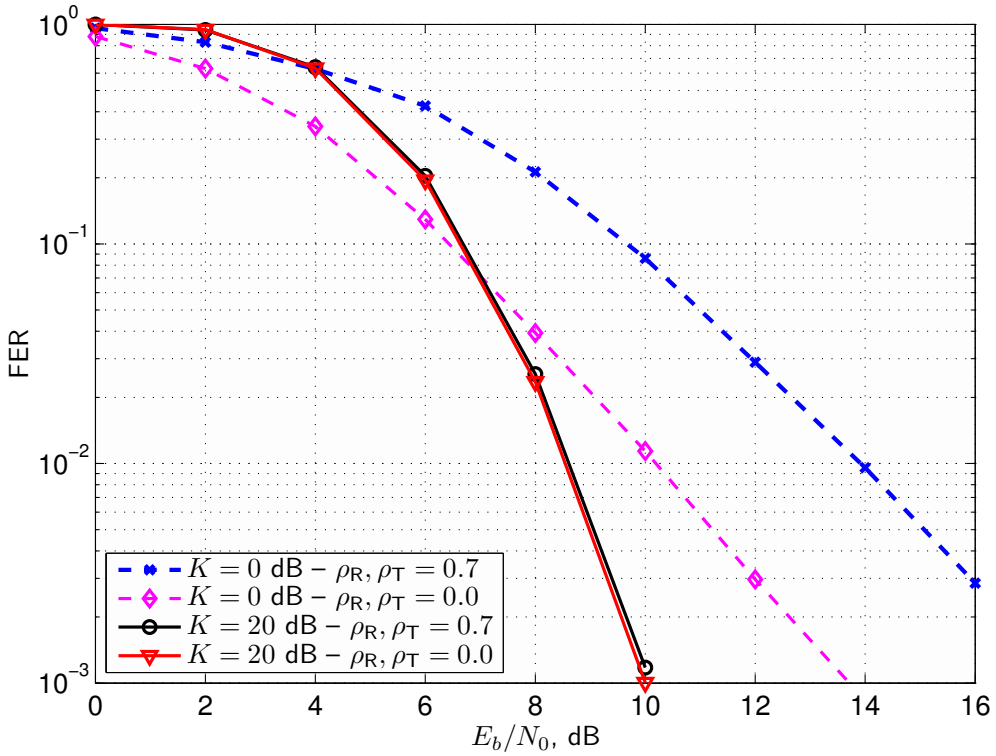


Figure 2.11. FER of a 2×2 Rician fading MIMO system versus E_b/N_0. Optimum receiver with $P = 4$

2.3.4 Correlated Rayleigh fading MIMO channel: Mismatched versus optimum receiver performance

Figure 2.12 illustrates the benefits of using the optimum receiver (versus the mismatched one) for a 2×2 MIMO system with Rayleigh fading and correlation coefficients $\rho_{\mathsf{T}} = \rho_{\mathsf{R}} = 0.7$ (i.e., the case of joint transmit and receive correlation).

The curves report the FERs corresponding to the minimum and optimum number of pilot symbol intervals per frame. The diagrams show that the optimum receiver gains approximately 1 dB when the *optimum* number of pilot symbol intervals per frame is used.

However, in order to achieve its best performance, the mismatched receiver needs to use a number of pilot symbol intervals per frame equal to 16, which reduces the throughput by approximately 11% whereas the throughput reduction entailed by the optimum receiver is only 3% for $P = 4$.

Fixing the throughput to its highest value (corresponding to setting $P = 2$), the optimum receiver gain increases to approximately 2.5 dB.

2.3.5 Correlated Rayleigh fading MIMO channel: Trained receiver

Figure 2.12 plots also the FER achieved by our trained receiver scheme implementing parameter estimation as described in Section 2.2.6 with a training phase of 10 blocks. The trained receiver FER overlaps almost exactly with the optimum receiver, which confirms that real parameter estimation (though obtained by a suboptimum algorithm) is not a source of performance degradation, at least in the case considered here. Other performance results, not reported here for space conciseness, show that this feature extends nicely to different MIMO channel scenarios so that it can safely stated that the proposed receiver is not substantially affected by real parameter estimation methods.

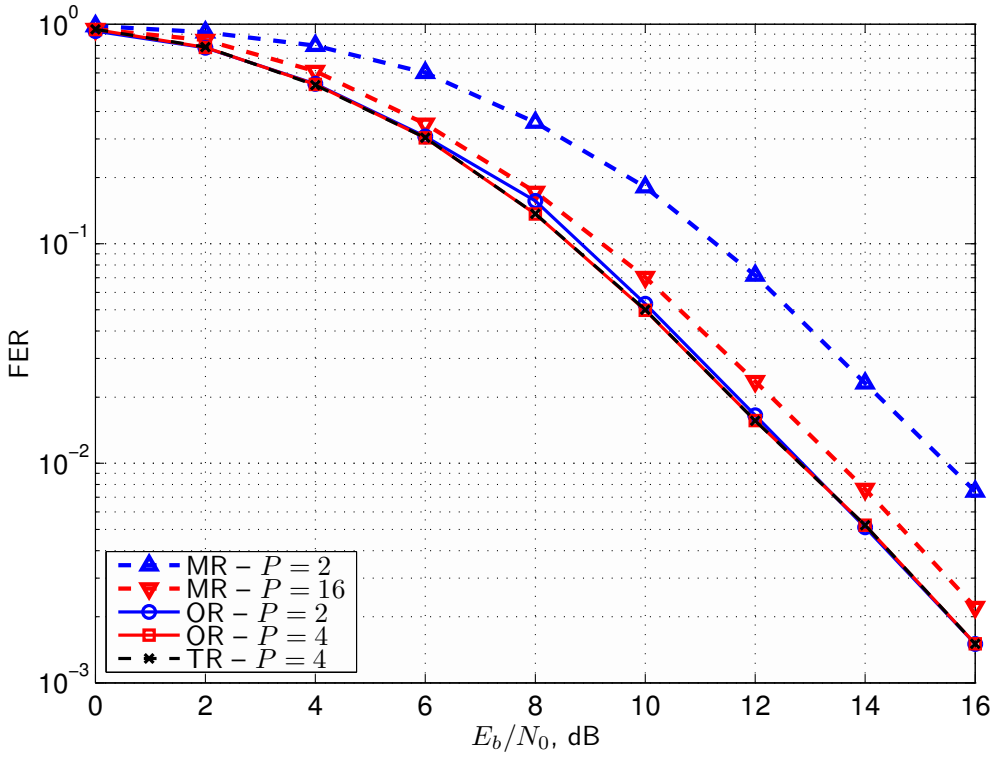


Figure 2.12. FER of a 2×2 Rayleigh fading MIMO system versus E_b/N_0. $\rho_{\mathsf{T}} = \rho_{\mathsf{R}} = 0.7$. Mismatched (MR) versus optimum receiver (OR) for different values of P. Trained receiver (TR) performance is also reported.

2.3.6 Additional receiver schemes

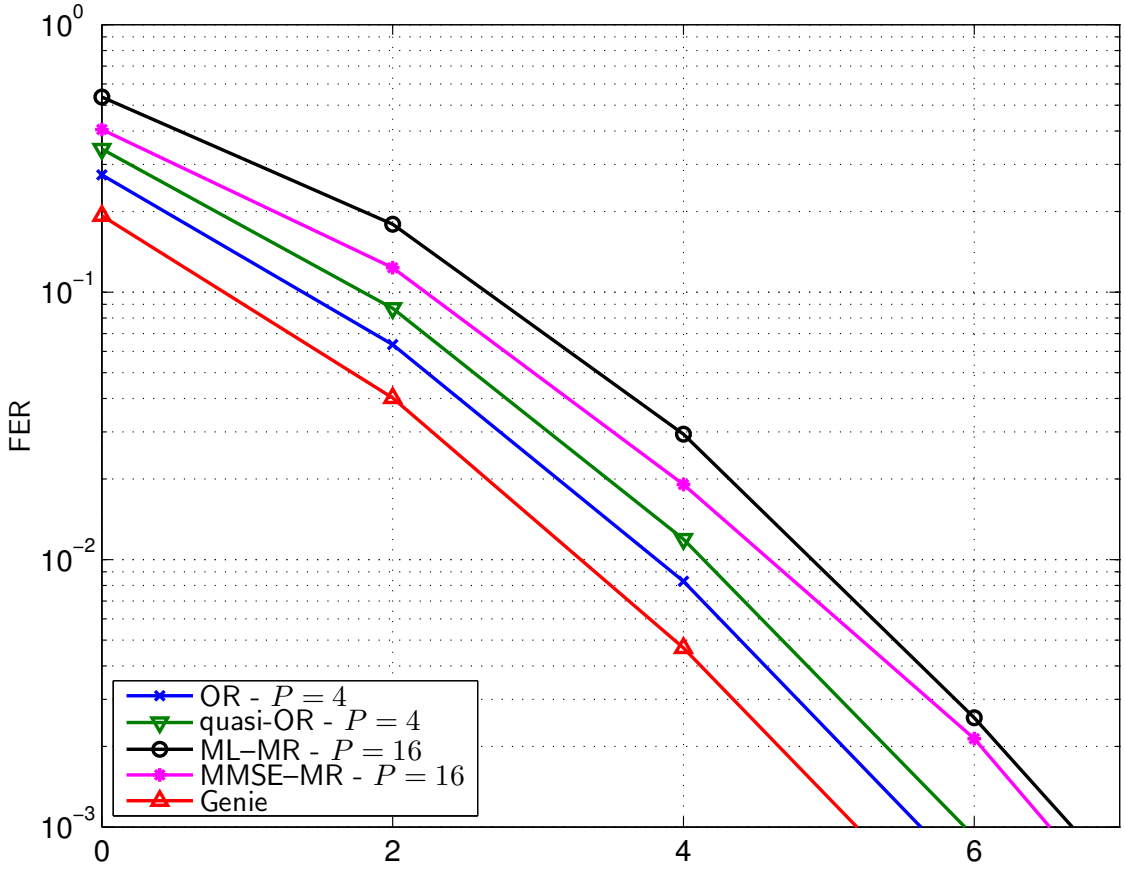


Figure 2.13. FER of a 2×8 Rician fading MIMO system versus E_b/N_0. $\rho_{\mathsf{T}} = \rho_{\mathsf{R}} = 0.7$ and $K = 10$ dB.

In this section the performance of two other receiver schemes, namely,

1. the MMSE-mismatched receiver described in Section 2.2.2;
2. the quasi-optimum receiver, obtained by assuming that the MIMO channel is uncorrelated and Rayleigh faded (as in [36]).

are considered

In order to emphasize the differences, a 2×8 Rician fading MIMO channel with $K = 10$ dB and $\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.7$ is considered. The FER versus E_b/N_0 results are reported in Figure 2.13. It can be seen that the receiver performance improves in the following order:

1. ML-mismatched,
2. MMSE-mismatched,
3. quasi-optimum,
4. optimum, and
5. genie-aided (ideal CSI).

2.4 Performance in a measured environment

2.4.1 Channel Measurements

In order to test the performance of our optimum receiver in a real scenario, the MIMO radio channel measurements obtained by *ftw* were used. The whole measurement data set is available from `http://www.ftw.at/measurements`
and fully documented in [41].

Two environments are considered:

1. The *urban* environment consists of an outdoor city scenario, with the receiver located on top of a building and the transmitter moving at pedestrian speed in the surroundings.

2. The *rich scattering* environment consists of an indoor scenario, with fixed receiver and transmitter located in two different rooms of the building. Moreover, around 35 pieces of tin foil have been distributed on the walls within the measurement area (see [41] for details).

The receive antenna array consists of a uniform linear array with 8 elements, spaced at a distance of $\lambda/2 = 7.5$ cm. The transmit antenna array consists of a moving uniform circular array (with diameter of 30 cm) with 15 elements spaced at $0.43\lambda = 6.45$ cm for the *urban* environment. The transmit array was uniform linear with 15 elements and $\lambda/2 = 7.5$ cm spacing for the *rich scattering* environment. Further details are available from [41].

Every data set consists of N_t time by N_f frequency samples of a 15×8 MIMO wideband transfer function with center frequency at 2 GHz.

For both environments two antenna settings have been defined

1. uncorrelated

2. correlated

by properly scheduling the sample extraction. The original data set was rearranged in order to obtain a data set for a 2×2 narrowband MIMO channel.

From each 8×15 time-frequency sample matrix, 28 2×2 matrices were obtained by fixing the antenna distance (7 transmit elements, 4 receive elements for the uncorrelated setting, 1 transmit/receive element for the correlated setting) and scanning the 8×15 channel matrix row-wise (see Figs. 2.14(a) and 2.14(b) for illustration).

Channel matrices were scanned first along the time dimension, then along the frequency dimension. From this sequence of $L_0 = 28 \cdot N_t \cdot N_f$ channel matrices of size 2×2 a smaller subset consisting of the first $N_{f0} < N_f$ sub-bands was selected

to limit the simulation time. Finally, to satisfy condition (2.3), channel realizations have been normalized by the constraint

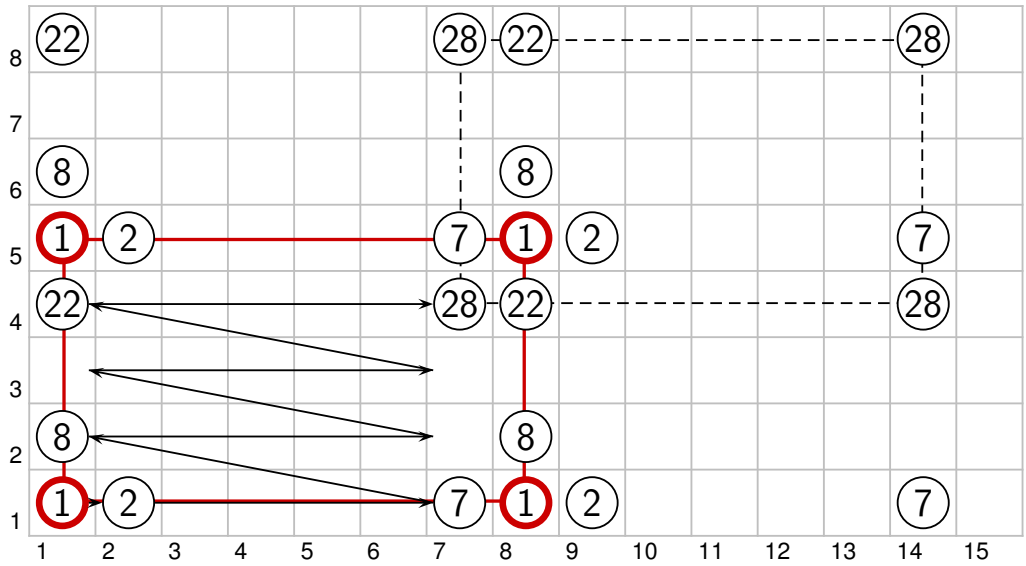
$$\frac{1}{L_0}\sum_{i=1}^{L_0}\mathbf{H}\mathbf{H}^{\mathsf{H}} = n_{\mathsf{T}}n_{\mathsf{R}} = 4.$$

2.4.2 Urban environment

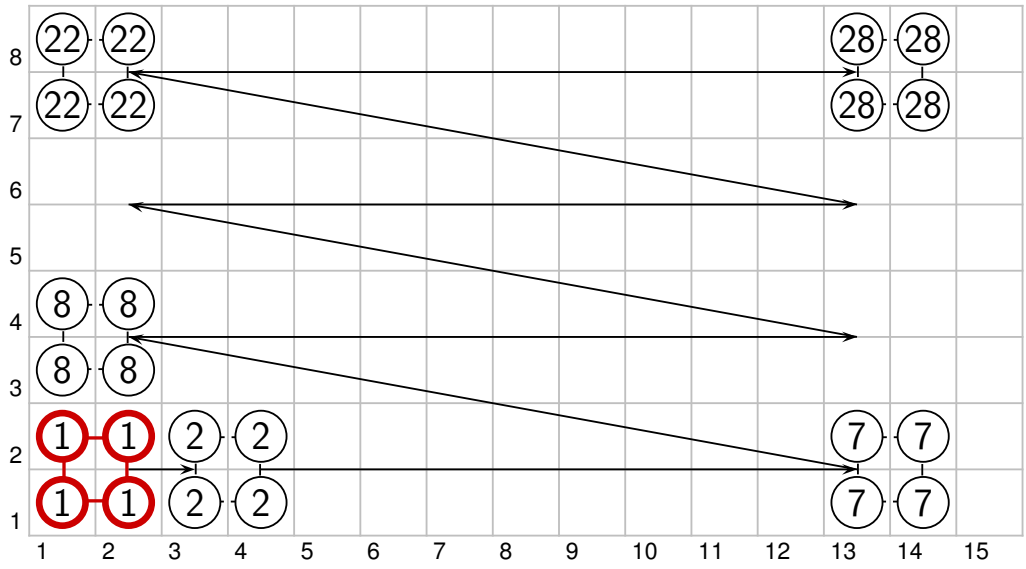
For the *urban* environment, the channel data set consists of $N_{f0} = 10$ sub-bands, each having $N_t = 311$ time samples with 28 2×2 matrices; as a result, $L_0 = 87080$.

Fig. 2.15 shows the FER performance in the uncorrelated case. It can be noticed that the optimum receiver loses approximately 0.4 dB from the genie receiver. Moreover, at the maximum throughput ($P = 2$), the mismatched receiver loses about 2.5 dB from the optimum and 3 dB from the genie receiver. This loss reduces at the price of increasing the number of pilot symbols per frame ($P = 16$) with a considerable throughput reduction.

Fig. 2.16 shows the FER performance in the correlated case (adjacent antennas). The relative receiver performances are similar to the uncorrelated case though the FER is worse in all cases because of increased correlation.



(a) Uncorrelated setting. Distance: 4 in a column, 7 in a row



(b) Correlated setting. Distance: 1 in a column, 1 in a row

Figure 2.14. 2×2 matrix construction from 8×15 matrix. Elements labelled with the same number belong to the same 2×2 matrix. The distance between elements is fixed. The matrix is scanned rowwise, from submatrx labelled 1 to matrix numbered 28. Note that column 15 is not scanned in order to avoid a double use of the same element

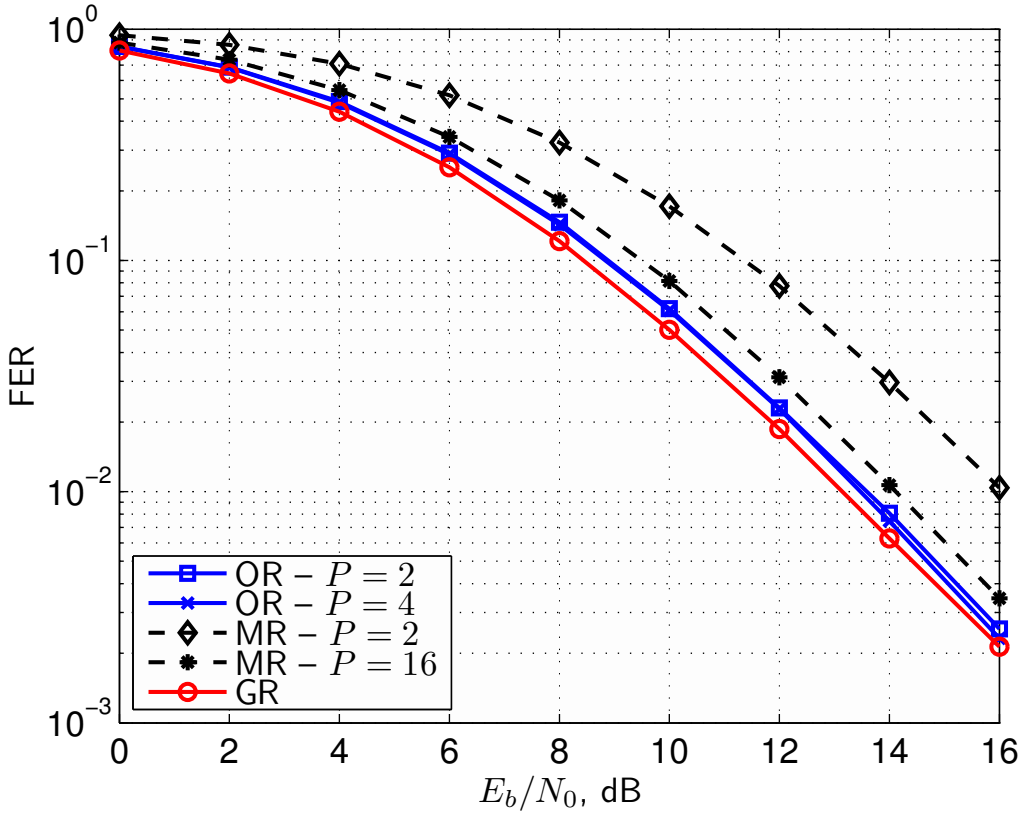


Figure 2.15. FER of a 2×2 MIMO system vs. E_b/N_0. Urban environment. $P = 2$ and $P = 4$. Mismatched and genie receiver FERs are also shown. Uncorrelated case.

2.4.3 Rich scattering environment

For the *rich scattering* environment, the channel data set consists of $N_{f0} = 20$ subbands, each having $N_t = 128$ time samples with 28 2×2 matrices; as a result, $L_0 = 71680$.

Fig. 2.17 shows the FER performance in the uncorrelated case. The figure shows that the optimum receiver loses about 0.5 dB from the genie receiver and that the mismatched receiver loses about 2.7 dB from the optimum receiver at maximum throughput ($P = 2$). It can then be noticed that the FER performance are slightly more spread in this case than in the urban environment.

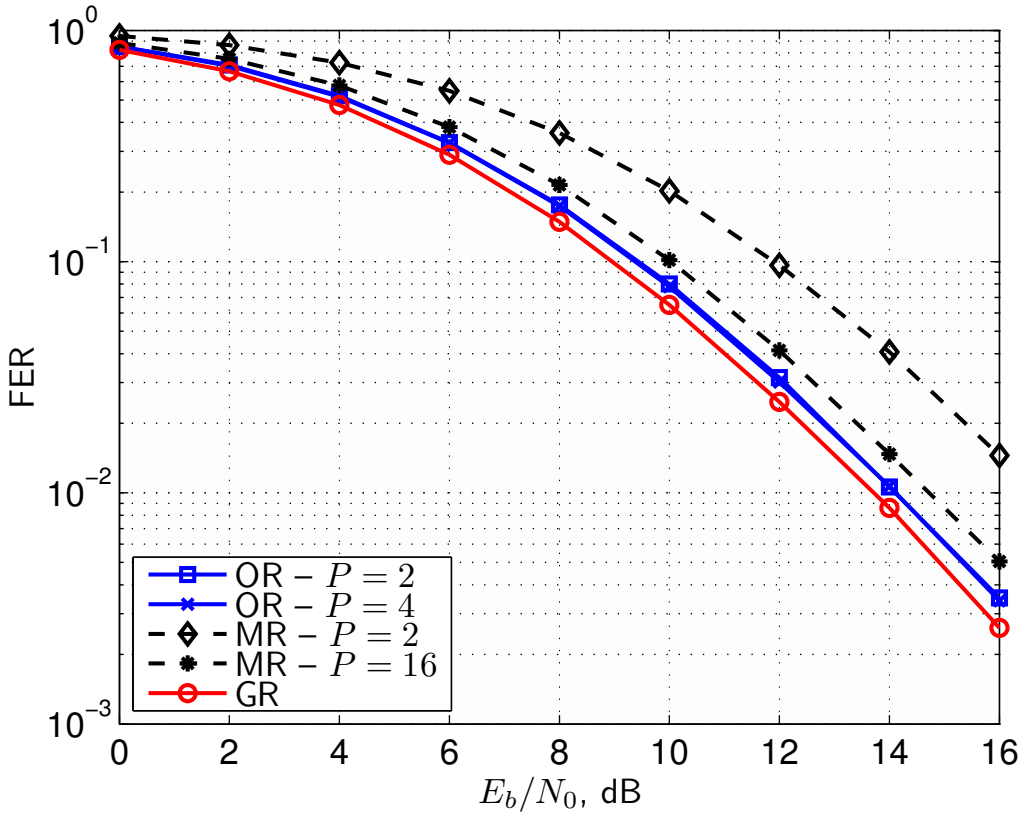


Figure 2.16. FER of a 2×2 MIMO system vs. E_b/N_0. Urban environment. $P = 2$ and $P = 4$. Mismatched and genie receiver FERs are also shown. Correlated case.

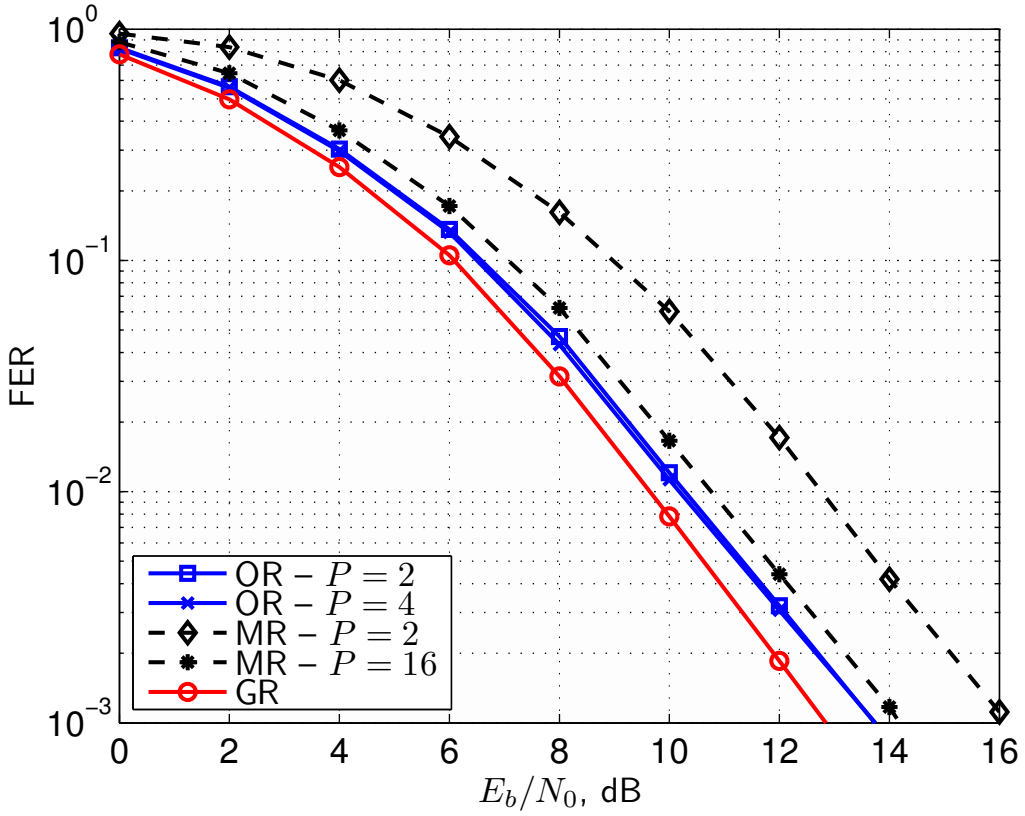


Figure 2.17. FER of a 2×2 MIMO system vs. E_b/N_0. Rich Scattering environment. $P = 2$ and $P = 4$. Mismatched and genie receiver FERs are also shown. Uncorrelated case.

2.4.4 Parameter estimation

Here the results of parameter estimation in the *rich scattering* environment for both uncorrelated and correlated settings are shown. The results are reported in Figs. 2.18, 2.19, and 2.20.

The Rice factor is estimated using (2.17) The receive and transmit correlation coefficients reported in Figs. 2.19 and 2.20 are defined as

$$\rho_{\mathsf{R}} = \frac{|(\widehat{\mathbf{R}})_{12}|}{\sqrt{(\widehat{\mathbf{R}})_{11}(\widehat{\mathbf{R}})_{22}}} \qquad \text{and} \qquad \rho_{\mathsf{T}} = \frac{|(\widehat{\mathbf{T}})_{12}|}{\sqrt{(\widehat{\mathbf{T}})_{11}(\widehat{\mathbf{T}})_{22}}} \tag{2.18}$$

The diagrams plot the behavior of the estimated parameters versus time at different values of E_b/N_0. Of course, in all cases, the most reliable parameters correspond to the highest value of E_b/N_0 (16 dB in our simulations).

It can be noticed that, at $E_b/N_0 = 16$ dB, the Rice factor estimation ranges from -14.7 dB (uncorrelated case) to -15.7 dB (correlated case). The receive correlation coefficient ranges between 0.2 and 0.5 in both cases. The transmit correlation coefficient shows instead (as expected) more noticeable differences in the correlated and uncorrelated cases: from 0.23 to 0.18 in the correlated case ($E_b/N_0 = 16$ dB); from 0.02 to 0.04 in the uncorrelated case.

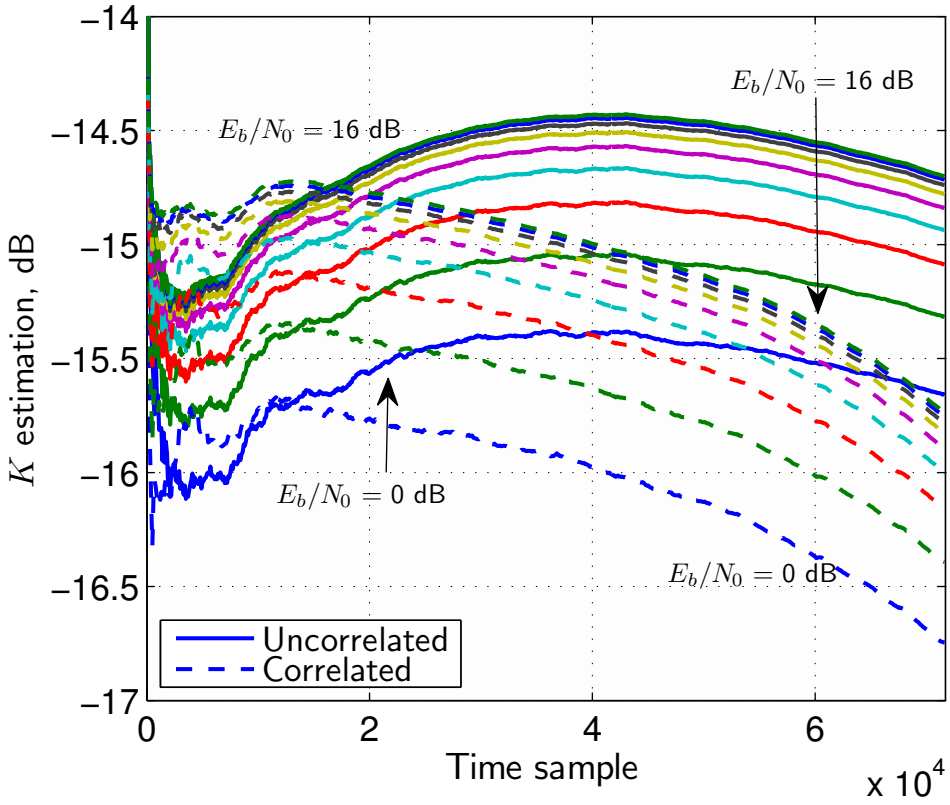


Figure 2.18. Rice Factor K estimation of a 2×2 MIMO system vs. time. Rich scattering environment. $P = 4$. Uncorrelated (solid) and correlated (dashed) case. E_b/N_0 dependence.

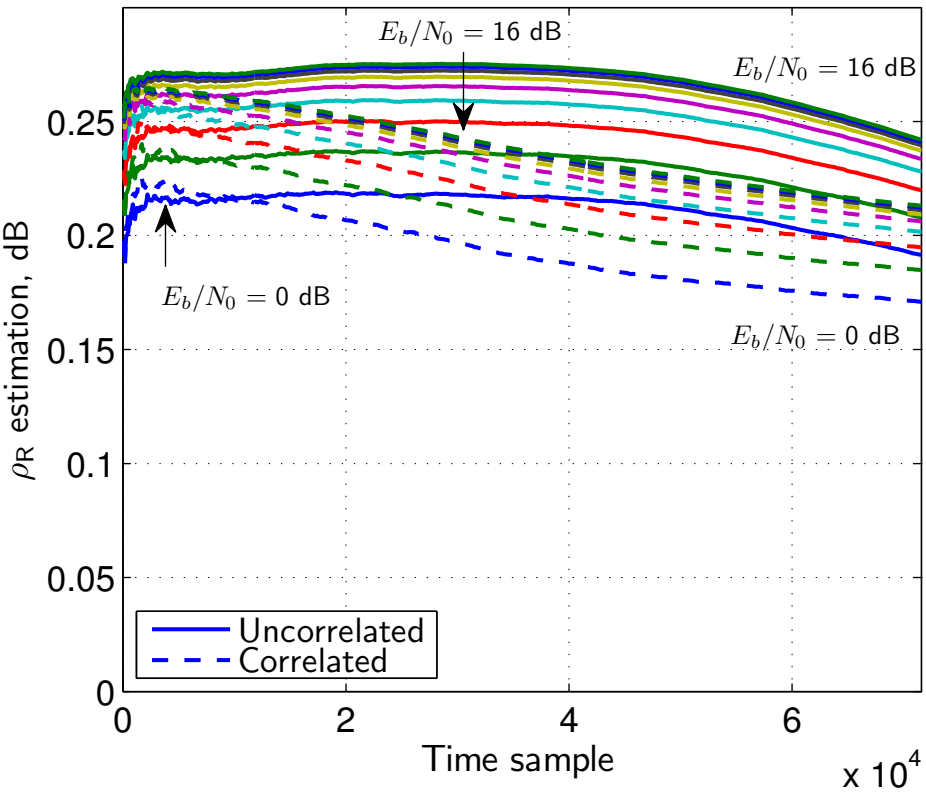


Figure 2.19. Receive correlation coefficient ρ_r estimation of a 2×2 MIMO system vs. time. Rich scattering environment. $P = 4$. Uncorrelated (solid) and correlated (dashed) case. E_b/N_0 dependence.

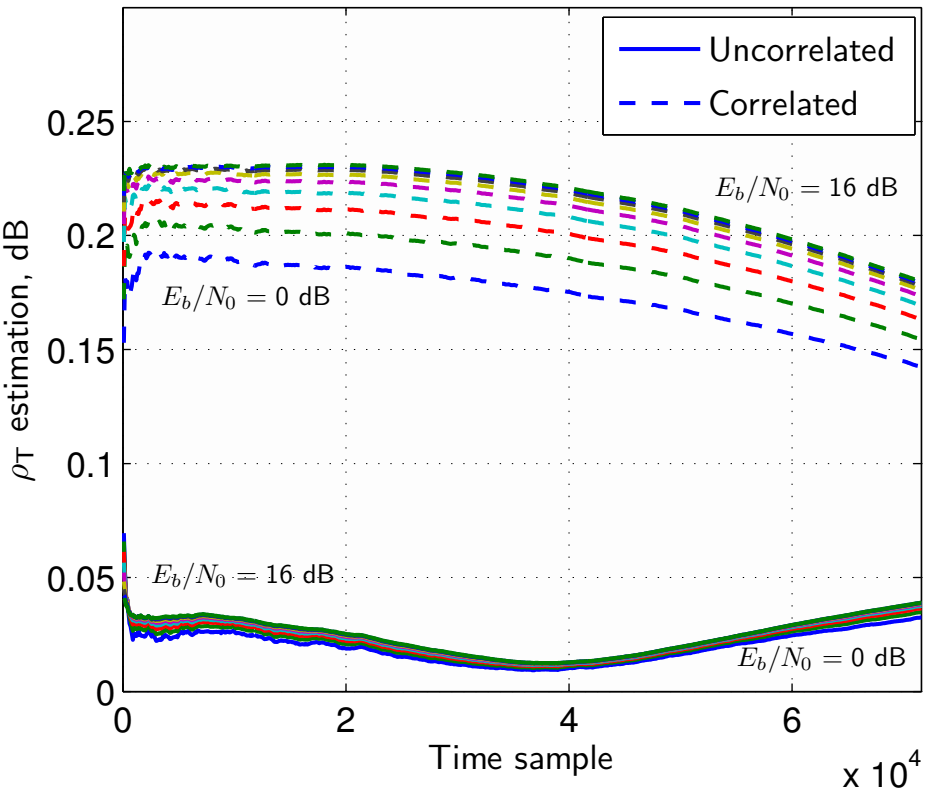


Figure 2.20. Transmit correlation coefficient ρ_t estimation of a 2×2 MIMO system vs. time. Rich scattering environment. $P = 4$. Uncorrelated (solid) and correlated (dashed) case. E_b/N_0 dependence.

2.5 Blind Optimum Receiver

2.5.1 System Model

A single-user narrowband MIMO block fading channel with n_{T} transmit and n_{R} receive antennas is considered. Each time frame consists of N symbol intervals of T time units. In a time frame, the transmitter sends an $n_{\mathsf{T}} \times N$ matrix $\mathbf{X}$ with average symbol energy

$$E_s \triangleq (n_{\mathsf{T}} N)^{-1} \mathbb{E}[\|\mathbf{X}\|^2] \ .$$

From the block fading assumption, the receiver observes the matrix $\mathbf{Y}$ given by:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{2.19}$$

where $\mathbf{Z}$, representing the receiver noise, has iid entries distributed as $\mathcal{N}_c(0,N_0)$.

The channel matrix is

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\mathbf{H}_0 + \sqrt{\frac{1}{K+1}}\mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2} \tag{2.20}$$

where

$$\bar{\mathbf{H}} \triangleq \mathbb{E}[\mathbf{H}] = \sqrt{\frac{K}{K+1}}\mathbf{H}_0$$

is the channel mean, $\mathbf{R}$ and $\mathbf{T}$ are the receive and transmit correlation matrices, respectively, and $\mathbf{W}$ is an $n_{\mathsf{R}} \times n_{\mathsf{T}}$ matrix with iid entries $\sim \mathcal{N}_c(0,1)$. The Rice factor is given by

$$K = \frac{\|\bar{\mathbf{H}}\|^2}{\mathbb{E}[\|\mathbf{H} - \bar{\mathbf{H}}\|^2]} \tag{2.21}$$

2.5.2 Receiver architecture

The blind receiver estimates the transmitted (equally likely) code words according to the rule

$$\widehat{\mathbf{X}} = \arg\max_{\mathbf{X}} \{\Pr(\mathbf{X}|\mathbf{Y})\} = \arg\max_{\mathbf{X}} \mathbb{E}_{\mathbf{H}}[p(\mathbf{Y}|\mathbf{X},\mathbf{H})] \tag{2.22}$$

which leads to the following decision metric associated to $\mathbf{X}$

$$\begin{aligned}\mu_{\mathsf{BR}}(\mathbf{X}) =& \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 + N_0 \ln\det(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}}} + \mathbf{C}(\mathbf{X})) \\ & - N_0 \mathrm{vec}\{\mathbf{B}(\mathbf{X})\}^{\mathsf{H}} (\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}}} + \mathbf{C}(\mathbf{X}))^{-1} \mathrm{vec}\{\mathbf{B}(\mathbf{X})\}\end{aligned} \tag{2.23}$$

where

$$\begin{cases} \mathbf{A}(\mathbf{X}) \triangleq \dfrac{\mathbf{T}^{1/2}(\mathbf{X}\mathbf{X}^{\mathsf{H}})\mathbf{T}^{1/2}}{N_0(K+1)} \\ \mathbf{B}(\mathbf{X}) \triangleq \dfrac{\mathbf{R}^{1/2}(\mathbf{Y}-\bar{\mathbf{H}}\mathbf{X})\mathbf{X}^{\mathsf{H}}\mathbf{T}^{1/2}}{N_0\sqrt{K+1}} \\ \mathbf{C}(\mathbf{X}) \triangleq \mathbf{A}(\mathbf{X})^{\mathsf{T}} \otimes \mathbf{R} \end{cases}$$

Remark 2.5.1. Note that this receiver requires perfect knowledge of the CDIR, i.e., of the fading statistic parameters $\bar{\mathbf{H}}$, K, $\mathbf{R}$ and $\mathbf{T}$. These can be estimated during an initial training phase, which is usually very short (see [51] for further details).

Iterative metric computation

The receiver metric can be calculated by an iterative algorithm. Define $\mathbf{X} = (\mathbf{X}^{-}, \mathbf{x})$ and $\mathbf{Y} = (\mathbf{Y}^{-}, \mathbf{y})$, where $\mathbf{X}^{-}$ and $\mathbf{Y}^{-}$ denote the *past* parts of the transmitted and received matrices and $\mathbf{x}$ and $\mathbf{y}$ represent the current symbol and received vector, respectively. Then, after some algebra, the metric (2.23) can be written as

$$\mu_{\mathsf{BR}}(\mathbf{X}) = \mu_{\mathsf{BR}}(\mathbf{X}^{-}) + \Delta\mu_{\mathsf{BR}}(\mathbf{X}^{-},\mathbf{x}) \ ,$$

where the metric increment is given by [51]:

$$\begin{aligned} \Delta\mu_{\mathsf{BR}}(\mathbf{X}^{-},\mathbf{x}) \triangleq & \|\mathbf{y}-\bar{\mathbf{H}}\mathbf{x}\|^2 \\ & + N_0 \ln\det\{\mathbf{I}_{n_{\mathsf{R}}} + \mathbf{\Psi}(\mathbf{x})^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X}^{-})\mathbf{\Psi}(\mathbf{x})\} \\ & - N_0 \mathrm{vec}\{\mathbf{B}(\mathbf{X})\}^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X})\mathrm{vec}\{\mathbf{B}(\mathbf{X})\} \\ & + N_0 \mathrm{vec}\left\{\mathbf{B}(\mathbf{X}^{-})\right\}^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X}^{-})\mathrm{vec}\left\{\mathbf{B}(\mathbf{X}^{-})\right\} \end{aligned} \tag{2.24}$$

where [51]:

$$\begin{cases} \mathbf{\Psi}(\mathbf{x}) & \triangleq \dfrac{(\mathbf{T}^{1/2}\mathbf{x})^{*}}{\sqrt{N_0(K+1)}} \otimes \mathbf{R}^{1/2} \\ \mathbf{\Lambda}(\mathbf{X}) & \triangleq \mathbf{\Lambda}(\mathbf{X}^{-}) + \Delta\mathbf{\Lambda}(\mathbf{X}^{-},\mathbf{x}) \\ \Delta\mathbf{\Lambda}(\mathbf{X}^{-},\mathbf{x}) & \triangleq -\mathbf{\Lambda}(\mathbf{X}^{-})\mathbf{\Psi}(\mathbf{x})[\mathbf{I}_{n_{\mathsf{R}}} + \mathbf{\Psi}(\mathbf{x})^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X}^{-})\mathbf{\Psi}(\mathbf{x})]^{-1}\mathbf{\Psi}(\mathbf{x})^{\mathsf{H}}\mathbf{\Lambda}(\mathbf{X}^{-}) \\ \mathbf{B}(\mathbf{X}) & \triangleq \mathbf{B}(\mathbf{X}^{-}) + \Delta\mathbf{B}(\mathbf{x}) \\ \Delta\mathbf{B}(\mathbf{x}) & \triangleq \mathbf{R}^{1/2}\dfrac{\mathbf{y}-\bar{\mathbf{H}}\mathbf{x}}{N_0\sqrt{K+1}}\mathbf{x}^{\mathsf{H}}\mathbf{T}^{1/2} \end{cases}$$

with

$$\mathbf{\Lambda}(\varnothing) = \mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}}$$

and

$$\mathbf{B}(\varnothing) = \mathbf{0}_{n_{\mathsf{T}}\times n_{\mathsf{R}}} \ .$$

Remark 2.5.2. The previous equations are derived directly from [51, Sec. III-C] by replacing the pilot matrix with an empty matrix. Simulation results show that the Viterbi decoder may experience numerical instability in the metric computation. This effect is eliminated by inserting an initial exhaustive decoding phase of n_{T} trellis steps at the beginning of each frame (code word). During this phase, the path metrics are computed directly, without using the iterative decoding algorithm.

Asymptotic SNR performance ($N_0 \to 0$)

Here it is assumed that the $\mathbf{X}\mathbf{X}^{\mathsf{H}}$ is invertible (i.e., the code word $\mathbf{X}$ has full row rank). Setting

$$\mathbf{A}_0(\mathbf{X}) \triangleq N_0 \mathbf{A}(\mathbf{X})$$

and

$$\mathbf{B}_0(\mathbf{X}) \triangleq N_0 \mathbf{B}(\mathbf{X}) \ ,$$

when $N_0 \to 0$ it can be obtained

$$\begin{aligned}
\mu_{\mathsf{BR}}(\mathbf{X}) =& \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 + N_0 \ln\det(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}}} + \mathbf{A}_0(\mathbf{X})^{\mathsf{T}} \otimes \mathbf{R}/N_0) \\
& - \mathrm{vec}\left\{\mathbf{B}_0(\mathbf{X})\right\}^{\mathsf{H}} (N_0 \mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}}} + \mathbf{A}_0(\mathbf{X})^{\mathsf{T}} \otimes \mathbf{R})^{-1} \mathrm{vec}\left\{\mathbf{B}_0(\mathbf{X})\right\} \\
=& \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 - \mathrm{Tr}\left(\mathbf{A}_0(\mathbf{X})^{-1} \mathbf{B}_0(\mathbf{X})^{\mathsf{H}} \mathbf{R}^{-1} \mathbf{B}_0(\mathbf{X})\right) \\
& - n_{\mathsf{T}} n_{\mathsf{R}} N_0 \ln N_0 + O(N_0) \\
=& \mathrm{Tr}\left((\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X})^{\mathsf{H}} (\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X})[\mathbf{I}_N - \mathbf{X}^{\mathsf{H}}(\mathbf{X}\mathbf{X}^{\mathsf{H}})^{-1}\mathbf{X}]\right) \\
& - n_{\mathsf{T}} n_{\mathsf{R}} N_0 \ln N_0 + O(N_0) \\
=& \mathrm{Tr}\left(\mathbf{Y}^{\mathsf{H}}\mathbf{Y}[\mathbf{I}_N - \mathbf{X}^{\mathsf{H}}(\mathbf{X}\mathbf{X}^{\mathsf{H}})^{-1}\mathbf{X}]\right) - n_{\mathsf{T}} n_{\mathsf{R}} N_0 \ln N_0 + O(N_0)
\end{aligned} \tag{2.25}$$

Neglecting the terms in (2.25) independent of $\mathbf{X}$, the GLRT receiver metric [47, eq. (34)] is obtained

$$\mu_{\mathsf{GLRT}}(\mathbf{X}) = -\mathrm{Tr}[\mathbf{Y}\mathbf{X}^{\mathsf{H}}(\mathbf{X}\mathbf{X}^{\mathsf{H}})^{-1}\mathbf{X}\mathbf{Y}^{\mathsf{H}}] \ .$$

Therefore, the GLRT receiver performance converges to that of the proposed blind receiver, as the SNR grows sufficiently large.

2.5.3 Numerical Results

Here the 4-state space–time code obtained from an optimum rate-2/4 binary convolutional code whose details are described in [36, Sec. VII-B] (STC-1) is considered.

Correlation matrices are exponential [11]: $(\mathbf{T})_{ij} = \rho_{\mathsf{T}}^{j-i}$ for $i \leq j$ and $(\mathbf{T})_{ji}^*$ otherwise. $(\mathbf{R})_{ij} = \rho_{\mathsf{R}}^{j-i}$ for $i \leq j$ and $(\mathbf{R})_{ji}^*$ otherwise. Following [59], set $(\mathbf{H}_0)_{ij} = 1$.

Our performance results are based on the effective throughput defined as [60]:

$$\eta \triangleq \frac{N}{N+P}\mu_b(1-\mathrm{FER}) \ , \tag{2.26}$$

where P is the number of pilot symbols per frame for the pilot-aided receiver, μ_b is the number of information bits per symbol, and FER is the frame error rate. The effective throughput η represents the average number of *error-free information bits* per symbol time transmitted with hybrid ARQ. Obviously, $\eta \leq \mu_b$.

It is assumed that the frame length (including data and pilot symbols) is $N+P = 24$, where $P = 0$ for the blind receiver (BR) and genie-aided receiver (GR) cases and 2,4 for the pilot-aided receiver (PR).

In order to account for the pilot power loss, it is assumed that the average bit energy is

$$E_b = \frac{(N+P)n_{\mathsf{T}}E_{\mathsf{S}}}{N\mu_b}$$

where $\mu_b = 2$ with the space–time code considered. Figs. 2.21 to 2.23 plot the effective throughput η versus E_b/N_0 for Rice factors $K = 0$,10, and 20 dB, respectively. Every figure refers to the uncorrelated ($\rho_{\mathsf{T}} = \rho_{\mathsf{R}} = 0$) and jointly correlated ($\rho_{\mathsf{T}} = \rho_{\mathsf{R}} = 0.7$) cases with four types of receivers:

1. the proposed BR,
2. the PR with $P = 2$,
3. the PR with $P = 4$ and
4. the GR.

It can be seen that the BR always outperforms the PR for $K = 10$ and 20 dB. However, in the uncorrelated case when $K = 0$ dB, the BR is optimum for E_b/N_0 above 6.4 dB, whereas the PR ($P = 2$) is optimum below that threshold. Moreover, the PR ($P = 4$) is slightly better than the PR ($P = 2$) when the E_b/N_0 falls below -3 dB.

Remark 2.5.3. Notice that the GLRT receiver is not included among the results because the code words of the space–time code considered do not always satisfy the requirement of having full row rank.

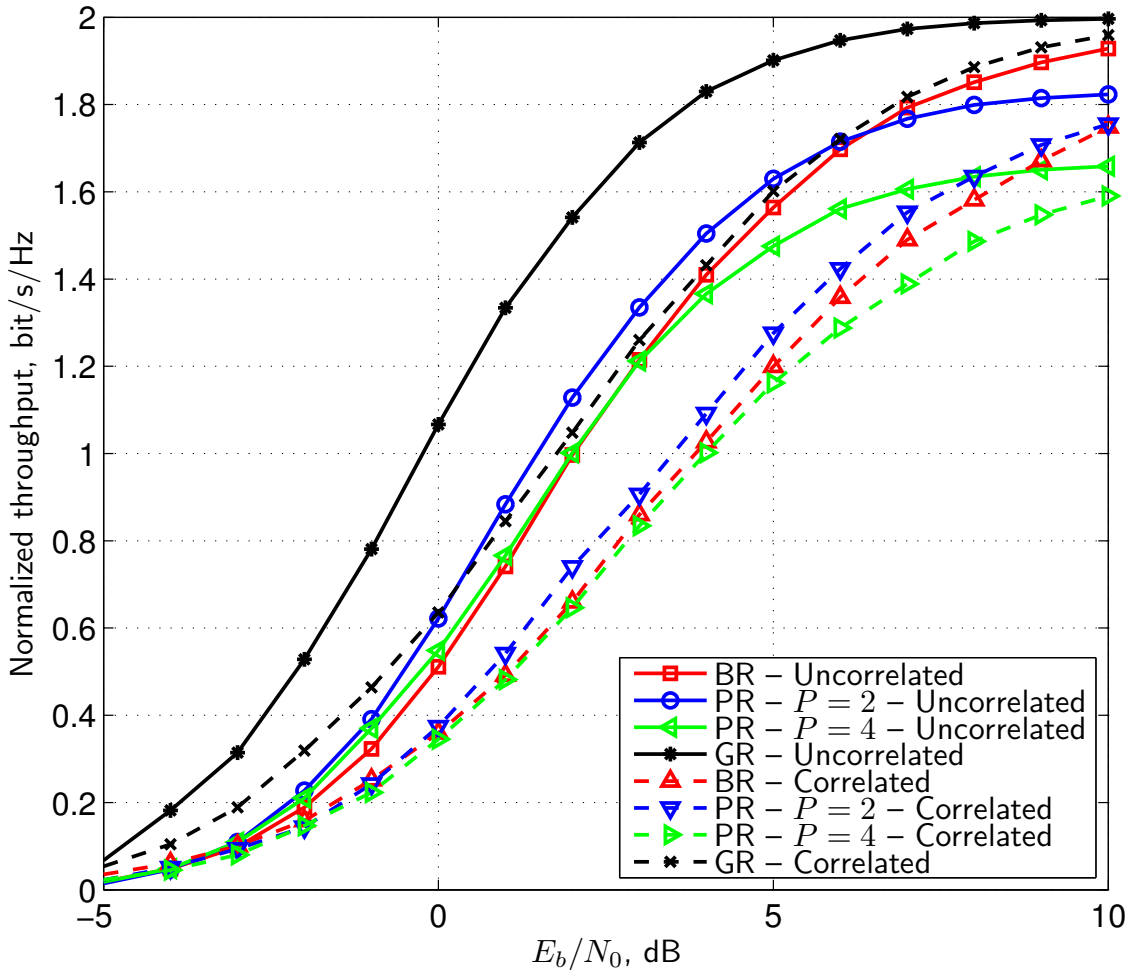


Figure 2.21. Normalized throughput η vs. E_b/N_0 of a 2×2 Rician fading MIMO channel with the blind, pilot-aided and genie receiver with and without correlation for $K = 0$ dB.

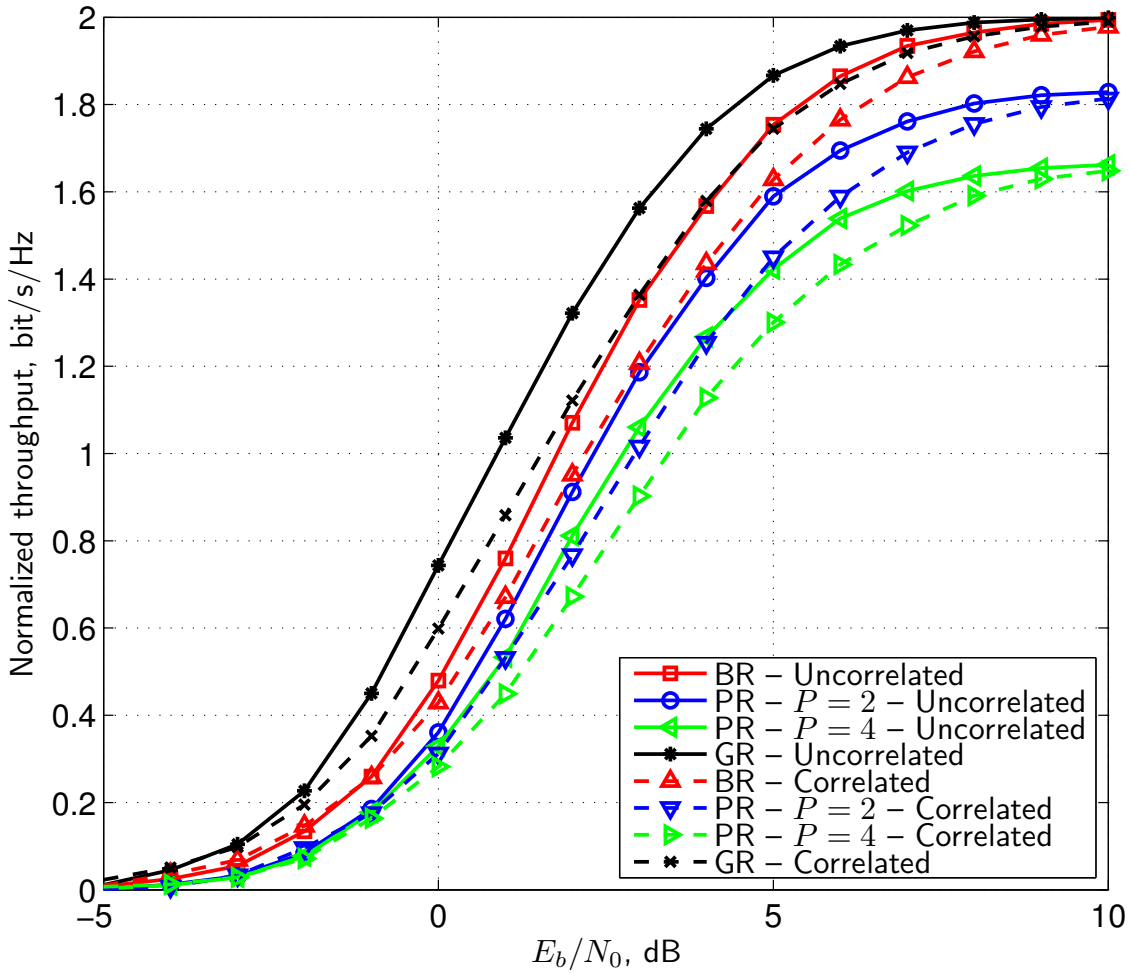


Figure 2.22. Normalized throughput η vs. E_b/N_0 of a 2×2 Rician fading MIMO channel with the blind, pilot-aided and genie receiver with and without correlation for $K = 10$ dB.

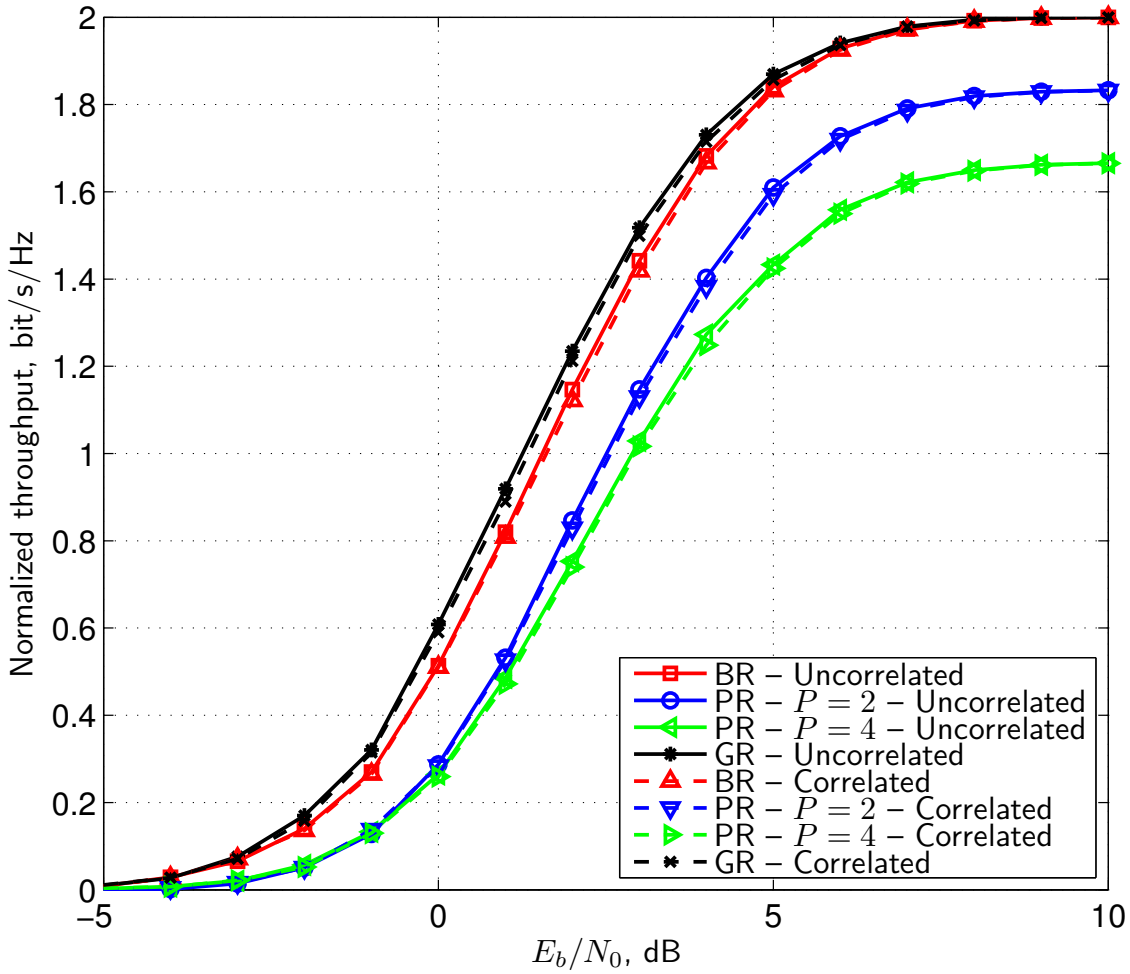


Figure 2.23. Normalized throughput η vs. E_b/N_0 of a 2×2 Rician fading MIMO channel with the blind, pilot-aided and genie receiver with and without correlation for $K = 20$ dB.

Chapter 3

Optimum MIMO-OFDM receiver

An optimum receiver for MIMO–OFDM communication systems accounting for realistic channel estimation is proposed.

The receiver is assumed to comply with the emerging IEEE 802.11n standard and its performance is compared against that of a *genie* receiver (corresponding to ideal channel estimation) and a *mismatched* receiver (using estimated channel state information in the ideal channel metric).

Receiver complexity is addressed in two steps. First, by developing an iterative expression of the receiver metric. Second, by implementing a spectral approximation, which allows a dramatic reduction of the receiver complexity with unnoticeable degradation.

Since the *optimum* receiver is based on the availability of channel distribution information, it is shown by numerical results that its estimation has a marginal effect on the error performance and does not represent an issue for the receiver implementation.

3.1 System Model

Consider a single-user wideband MIMO-OFDM block fading channel with n_{F} subcarriers, n_{T} transmit and n_{R} receive antennas described by the following equation:

$$\tilde{\mathbf{y}}_n = \sum_{l=0}^{L-1} \widetilde{\mathbf{H}}_l \tilde{\mathbf{x}}_{n-l} + \tilde{\mathbf{z}}_n \qquad n = -L+1, \ldots, n_{\mathsf{F}} \ , \tag{3.1}$$

where n is the discrete-time index, $\tilde{\mathbf{x}} \in \mathbb{C}^{n_{\mathsf{T}} \times 1}$ is the column vector of information symbols transmitted at time n, $\widetilde{\mathbf{H}}_l \in \mathbb{C}^{n_{\mathsf{R}} \times n_{\mathsf{T}}}$ is a sequence of matrices describing the channel impulse response, $\tilde{\mathbf{z}}_n \in \mathbb{C}^{n_{\mathsf{R}} \times 1}$ is the noise vector at time n, with $\tilde{\mathbf{z}}_n \sim \mathcal{N}_c(\mathbf{0}, N_0 \mathbf{I}_{n_{\mathsf{R}}})$, and $\tilde{\mathbf{y}}_n \in \mathbb{C}^{n_{\mathsf{R}} \times 1}$ is the received signal vector at time n.

According to the standard OFDM approach, consider a block of $(L + n_{\mathsf{F}})$ consecutive time intervals and assume that the transmitted symbol vectors satisfy the cyclic prefix condition, namely,

$$\tilde{\mathbf{x}}_n = \tilde{\mathbf{x}}_{n+n_{\mathsf{F}}} \ ,$$

for $n = -L+1, \ldots, 0$.

Applying a Discrete Fourier Transform (DFT) of size n_{F} to the block of received vectors in (3.1) yields the following frequency-domain equation:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k \qquad k = 1, \ldots, n_{\mathsf{F}} \ , \tag{3.2}$$

where k is the subcarrier index and (setting $\omega = \exp(-j2\pi/n_{\mathsf{F}})$) the following relations are defined

$$\begin{cases} \mathbf{x}_k = \sum_{n=1}^{n_{\mathsf{F}}} \tilde{\mathbf{x}}_n \omega^{(k-1)(n-1)} \\ \mathbf{y}_k = \sum_{n=1}^{n_{\mathsf{F}}} \tilde{\mathbf{y}}_n \omega^{(k-1)(n-1)} \\ \mathbf{z}_k = \sum_{n=1}^{n_{\mathsf{F}}} \tilde{\mathbf{z}}_n \omega^{(k-1)(n-1)} \\ \mathbf{H}_k = \sum_{n=1}^{n_{\mathsf{F}}} \tilde{\mathbf{H}}_n \omega^{(k-1)(n-1)} \end{cases} .$$

Here, $\mathbf{H}_k$ is the channel frequency response sampled at the k-th subcarrier frequency. Assume that the fading process is sufficiently slow to consider it constant for the duration of a time frame (*quasi-static* assumption). Each frame is composed by P pilot and N data symbol intervals.

Following a standard approach, the pilot and information symbols are generated in the frequency domain, then processed with a set of inverse DFT (IDFT) and

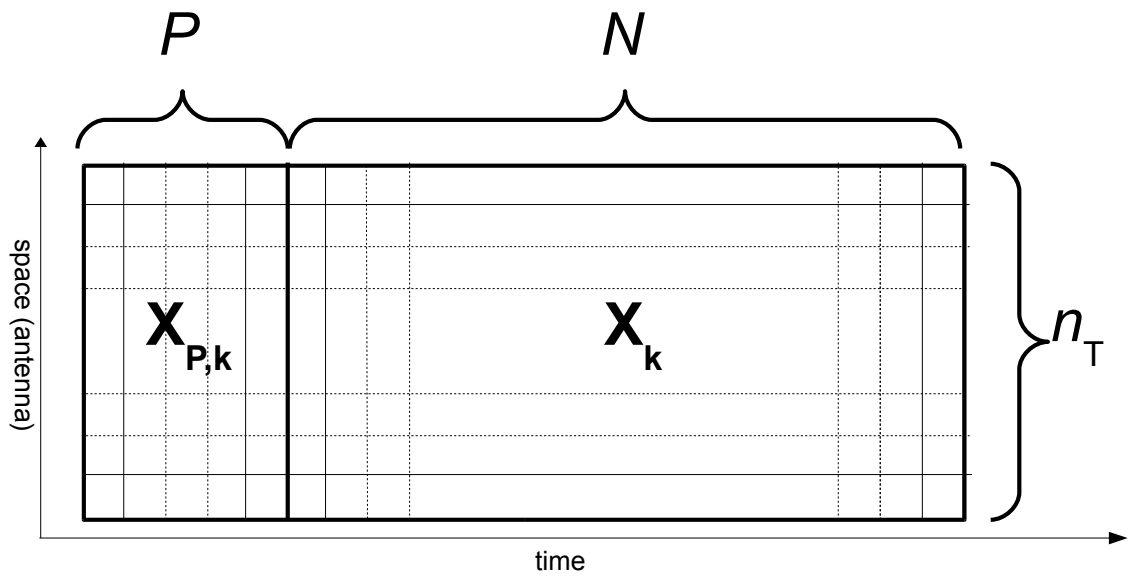


Figure 3.1. The subcarrier k frame structure

then transmitted over the wireless channel. Thus, during each frame and for each subcarrier, the transmitter sends a matrix $\mathbf{X}_{\mathsf{P},k} \in \mathbb{C}^{n_\mathsf{T} \times P}$ of pilot symbols followed by a matrix $\mathbf{X}_k \in \mathbb{C}^{n_\mathsf{T} \times N}$ of data symbols, as depicted in Figure 3.1.

The pilot symbol matrices $\mathbf{X}_{\mathsf{P},k}$ are assumed to have full-rank, so that $P \geq n_\mathsf{T}$. Collecting the symbols from the same time frame, the channel equation corresponding to one time frame can be written in the frequency domain as follows:

$$\mathbf{Y}_{\mathsf{P},k} = \mathbf{H}_k \mathbf{X}_{\mathsf{P},k} + \mathbf{Z}_{\mathsf{P},k}, \qquad \mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{Z}_k \ , \tag{3.3}$$

The noise affecting different subcarriers is assumed to be uncorrelated, in accordance with [61, Sec. II.B].

Next, defining

$$\begin{cases} \mathbf{X}_\mathsf{P} \triangleq (\mathbf{X}_{\mathsf{P},1}^\mathsf{T}, \ldots, \mathbf{X}_{\mathsf{P},n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \\ \mathbf{Y}_\mathsf{P} \triangleq (\mathbf{Y}_{\mathsf{P},1}^\mathsf{T}, \ldots, \mathbf{Y}_{\mathsf{P},n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \\ \mathbf{Z}_\mathsf{P} \triangleq (\mathbf{Z}_{\mathsf{P},1}^\mathsf{T}, \ldots, \mathbf{Z}_{\mathsf{P},n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \end{cases}$$

and

$$\begin{cases} \mathbf{X} \triangleq (\mathbf{X}_1^\mathsf{T}, \ldots, \mathbf{X}_{n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \\ \mathbf{Y} \triangleq (\mathbf{Y}_1^\mathsf{T}, \ldots, \mathbf{Y}_{n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \\ \mathbf{Z} \triangleq (\mathbf{Z}_1^\mathsf{T}, \ldots, \mathbf{Z}_{n_\mathsf{F}}^\mathsf{T})^\mathsf{T} \end{cases}$$

the subcarrier equations can be merged as follows:

$$\mathbf{Y}_\mathsf{P} = \mathbf{H}\mathbf{X}_\mathsf{P} + \mathbf{Z}_\mathsf{P}, \qquad \mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \ , \tag{3.4}$$

where the channel matrix $\mathbf{H}$ is block diagonal and defined as:

$$\mathbf{H} \triangleq \operatorname{diag}\big(\mathbf{H}_1,\ldots,\mathbf{H}_{n_\mathsf{F}}\big) = \bigoplus_{k=1}^{n_\mathsf{F}} \mathbf{H}_k \ . \tag{3.5}$$

3.1.1 Channel model

The covariance structure of $\mathbf{H}$ is modeled by extending the narrowband Kronecker model [20] to the wideband case after including the frequency domain as follows:

$$\operatorname{cov}\left((\mathbf{H}_k)_{ij},(\mathbf{H}_{k'})_{i'j'}\right) = (\mathbf{R})_{ii'}(\mathbf{T})^*_{jj'}(\mathbf{F})_{kk'} \tag{3.6}$$

where

$$\operatorname{cov}(X,Y) \triangleq \mathbb{E}\left[XY^*\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]^* \ .$$

The matrix $\mathbf{R} \in \mathbb{C}^{n_\mathsf{R}\times n_\mathsf{R}}$ and the matrix $\mathbf{T} \in \mathbb{C}^{n_\mathsf{T}\times n_\mathsf{T}}$ are the receive and transmit correlation matrices, respectively, and the matrix $\mathbf{F} \in \mathbb{C}^{n_\mathsf{F}\times n_\mathsf{F}}$ is the subcarrier correlation matrix. Note that all matrices $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{F}$ are positive semidefinite, $i,i' = 1,\ldots,n_\mathsf{R}$, $j,j' = 1,\ldots,n_\mathsf{T}$ and $k,k' = 1,\ldots,n_\mathsf{F}$.

From the definition (3.6), the matrix expression of $\mathbf{H}_k$ can be written as follows:

$$\mathbf{H}_k = \bar{\mathbf{H}}_k + \left((\mathbf{F}^{1/2})_k \otimes \mathbf{R}^{1/2}\right)\mathbf{W}\mathbf{T}^{1/2} \tag{3.7}$$

where $\bar{\mathbf{H}}_k \triangleq \mathbb{E}\left[\mathbf{H}_k\right]$. Here, $\mathbf{W} \in \mathbb{C}^{n_\mathsf{R}n_\mathsf{F}\times n_\mathsf{T}}$ is a matrix containing iid $\mathcal{N}_c(0,1)$-distributed entries. The equivalence between (3.6) and (3.7) is stated in the following lemma.

Lemma 3.1.1. The definition of (3.7) satisfies the covariance condition (3.6).

Proof. See Appendix B.1. □

Since $\bar{\mathbf{H}}_k$ is the matrix of LOS components corresponding to the kth subcarrier, the Rice factor for each carrier k can be defined in accordance with [61]:

$$K_k \triangleq \frac{\|\bar{\mathbf{H}}_k\|^2}{\mathbb{E}\left[\|\mathbf{H}_k - \bar{\mathbf{H}}_k\|^2\right]}.$$

According to the previous definition the following normalization holds:

$$\|\bar{\mathbf{H}}_k\|^2 = K_k(\mathbf{F})_{kk}\operatorname{Tr}\left(\mathbf{T}\right)\operatorname{Tr}\left(\mathbf{R}\right).$$

Define also $\bar{\mathbf{H}} \triangleq \mathbb{E}\left[\mathbf{H}\right] = \bigoplus_{k=1}^{n_\mathsf{F}} \bar{\mathbf{H}}_k$.

Assuming further

1. iid power allocation in space and time and
2. the same average symbol energy E_{S} and average pilot energy E_{P} for each carrier k, i.e.

$$\mathbb{E}\left[\mathbf{X}_k\mathbf{X}_k^{\mathsf{H}}\right] = NE_{\mathsf{S}}\mathbf{I}_{n_{\mathsf{T}}}, \qquad \mathbb{E}\left[\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}\right] = PE_{\mathsf{P}}\mathbf{I}_{n_{\mathsf{T}}},$$

the received data signal to noise ratio for each carrier can be written as:

$$\mathsf{SNR}_k = \frac{E_{\mathsf{S}}\mathrm{Tr}\left(\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{T}\right)\left(\mathbf{F}\right)_{kk}(K_k+1)}{N_0 n_{\mathsf{R}}} \tag{3.8}$$

and the received pilot signal to noise ratio for each carrier as:

$$\mathsf{SNR}_{\mathsf{P},k} = \frac{E_{\mathsf{P}}\mathrm{Tr}\left(\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{T}\right)\left(\mathbf{F}\right)_{kk}(K_k+1)}{N_0 n_{\mathsf{R}}}\ . \tag{3.9}$$

3.1.2 Estimation of the CDIR

According to a well known paradigm [62], fading occurs over different time scales. In our framework, short term variations are captured by the channel matrix $\mathbf{W}$, which changes from one frame to another; long term variations are accounted for by the channel matrix parameters $\bar{\mathbf{H}}$, $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{F}$, which are supposed to remain constant for a sufficiently large number of time frames [18].

In the following first the ideal case of perfectly known channel parameters at the receiver (i.e., perfect CDIR) is considered. Next, the more realistic case where the receiver estimates the relevant CDIR by using sample averages during a preliminary training phase (Section 3.2.7) is addressed.

3.2 Receiver architecture

In this section receiver designs which are not based on the common assumption of perfect CSIR are considered. The problem of CSIR recovery by the means of pilot symbols insertion [27] is also addressed.

Three different receiver architectures are considered:

1. The *genie Receiver*, which is an ideal receiver obtaining perfect CSIR from a *genie* and representing an error performance lower bound equivalent to the perfect CSIR assumption.

2. The *mismatched receiver*, which first estimates the CSIR, then uses it in a perfect CSIR metric.

3. The *optimum receiver*, which does not estimate the CSIR, but performs *joint* channel estimation and decoding.

In all cases, perfect CDIR is assumed.

Next, the optimum receiver performance will be studied by implementing an explicit CDIR estimation scheme. Channel coding is performed across subcarriers and a sequential receiver architecture suitable to Viterbi decoding is derived. In the following explicit descriptions of each receiver is given.

3.2.1 Genie Receiver

The ideal genie-aided receiver (yielding a lower bound to the FER) output is given by:

$$\widehat{\mathbf{X}}_{\mathsf{GR}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{GR}}(\mathbf{X}|\mathbf{Y}) \ , \tag{3.10}$$

where

$$\mu_{\mathsf{GR}}(\mathbf{X}|\mathbf{Y}) = \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 = \sum_{k=1}^{n_{\mathsf{F}}} \|\mathbf{Y}_k - \mathbf{H}_k\mathbf{X}_k\|^2 \tag{3.11}$$

3.2.2 Mismatched Receiver

Here the approach proposed in [36, 51] is extended to obtain a maximum-likelihood (ML) estimation of $\mathbf{H}_k$:

$$\widehat{\mathbf{H}}_k = \mathbf{Y}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}(\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}})^{-1} \tag{3.12}$$

provided that $(\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}})$ is invertible and hence the number of pilot intervals must satisfy the inequality $P \geq n_{\mathsf{T}}$, which is assumed to hold throughout the paper.

Remark 3.2.1. It is worth noting that the mismatched receiver does not require CDIR (i.e., the channel parameters $\bar{\mathbf{H}}$, $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{F}$), contrary to the optimum receiver.

A minimum-mean-square-error (MMSE) estimation can be also obtained. In section 2.2.2 it is shown that ML and MMSE estimates tend to each other when SNR grows large. Moreover, it is shown that the MMSE estimate requires the knowledge of CDIR parameters.

After obtaining the channel matrix estimates $\widehat{\mathbf{H}}_k$, the second stage of the mismatched receiver uses them as if they were exact, and outputs the code word

$$\widehat{\mathbf{X}}_{\mathsf{MR}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{MR}}(\mathbf{X}|\mathbf{Y}) \tag{3.13}$$

where

$$\mu_{\mathsf{MR}}(\mathbf{X}|\mathbf{Y}) = \sum_{k=1}^{n_{\mathsf{F}}} \|\mathbf{Y}_k - \widehat{\mathbf{H}}_k \mathbf{X}_k\|^2 \tag{3.14}$$

3.2.3 Optimum Receiver

Contrary to the mismatched receiver, the optimum receiver is not based on the explicit estimation of the channel matrix. Instead, it decodes the transmitted code word $\mathbf{X}$ by jointly processing the received signal matrices $\mathbf{Y}_{\mathsf{P}}$,$\mathbf{Y}$ and the pilot symbol matrix $\mathbf{X}_{\mathsf{P}}$. The resulting receiver output is given by

$$\widehat{\mathbf{X}}_{\mathsf{OR}} = \arg\min_{\mathbf{X}} \mu_{\mathsf{OR}}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_{\mathsf{P}},\mathbf{X}_{\mathsf{P}}) \tag{3.15}$$

where

$$\mu_{\mathsf{OR}}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_{\mathsf{P}},\mathbf{X}_{\mathsf{P}}) = -\ln \Pr\left(\mathbf{X}|\mathbf{X}_p,\mathbf{Y},\mathbf{Y}_p\right) \tag{3.16}$$

Assuming equally likely transmitted code words and the conditional independence of $\mathbf{Y}$ and $\mathbf{Y}_{\mathsf{P}}$ given $\mathbf{X}$,$\mathbf{X}_{\mathsf{P}}$, and $\mathbf{H}$, respectively, the probability in (3.16) can be written as

$$\begin{aligned}
\Pr\left(\mathbf{X}|\mathbf{X}_{\mathsf{P}},\mathbf{Y},\mathbf{Y}_{\mathsf{P}}\right) =& \frac{p\left(\mathbf{X},\mathbf{Y}\right) p\left(\mathbf{X}_{\mathsf{P}},\mathbf{Y}_{\mathsf{P}}\right)}{p\left(\mathbf{X}_{\mathsf{P}},\mathbf{Y},\mathbf{Y}_{\mathsf{P}}\right)} \\
=& \frac{\Pr\left(\mathbf{X}\right)\Pr\left(\mathbf{X}_{\mathsf{P}}\right)}{p\left(\mathbf{X}_{\mathsf{P}},\mathbf{Y},\mathbf{Y}_{\mathsf{P}}\right)} p\left(\mathbf{Y},\mathbf{Y}_{\mathsf{P}}|\mathbf{X},\mathbf{X}_{\mathsf{P}}\right) \\
=& \mathbb{E}_{\mathbf{H}}\left[p\left(\mathbf{Y},\mathbf{Y}_{\mathsf{P}}|\mathbf{X},\mathbf{X}_{\mathsf{P}},\mathbf{H}\right)\right] \\
=& \mathbb{E}_{\mathbf{H}}\left[p\left(\mathbf{Y}|\mathbf{X},\mathbf{H}\right) p\left(\mathbf{Y}_{\mathsf{P}}|\mathbf{X}_{\mathsf{P}},\mathbf{H}\right)\right] \ .
\end{aligned} \tag{3.17}$$

Then, the a posteriori probability in equation (3.17) can be calculated by writing explicitly the probability density functions involved:

Proposition 3.2.1. The a posteriori probability (3.17) can be evaluated explicitly as follows:

$$\begin{aligned}\Pr\left(\mathbf{X}|\mathbf{X}_p,\mathbf{Y},\mathbf{Y}_p\right) =&(\pi N_0)^{-(N+P)n_{\mathsf{R}}n_{\mathsf{F}}}\exp\left(-\frac{\|\mathbf{Y}_{\mathsf{P}}-\bar{\mathbf{H}}\mathbf{X}_{\mathsf{P}}\|^2+\|\mathbf{Y}-\bar{\mathbf{H}}\mathbf{X}\|^2}{N_0}\right)\\ &\cdot\det(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\mathbf{C})^{-1}\exp\left(\mathbf{b}^{\mathsf{H}}(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\mathbf{C})^{-1}\mathbf{b}\right)\end{aligned}\tag{3.18}$$

where the following definitions hold:

$$\begin{cases}\mathbf{b} & \triangleq & \displaystyle\sum_{k=1}^{n_{\mathsf{F}}}\mathrm{vec}\left\{(\mathbf{F}^{1/2})_k^{\mathsf{H}}\otimes\mathbf{R}^{1/2}\mathbf{B}_k\right\}\\ \mathbf{B}_k & \triangleq & \dfrac{(\mathbf{Y}_k-\bar{\mathbf{H}}_k\mathbf{X}_k)\mathbf{X}_k^{\mathsf{H}}+(\mathbf{Y}_{\mathsf{P},k}-\bar{\mathbf{H}}_k\mathbf{X}_{\mathsf{P},k})\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}}{N_0}\mathbf{T}^{1/2}\\ \mathbf{C} & \triangleq & \displaystyle\sum_{k=1}^{n_{\mathsf{F}}}\mathbf{A}_k^{\mathsf{T}}\otimes((\mathbf{F}^{1/2})_k^{\mathsf{H}}(\mathbf{F}^{1/2})_k)\otimes\mathbf{R}\\ \mathbf{A}_k & \triangleq & \dfrac{\mathbf{T}^{1/2}(\mathbf{X}_k\mathbf{X}_k^{\mathsf{H}}+\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}})\mathbf{T}^{1/2}}{N_0}.\end{cases}$$

Hence, after proper scaling and dropping terms independent of $\mathbf{X}$, the decision metric for the optimum receiver is obtained as:

$$\begin{aligned}\mu_{\mathsf{OR}}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_{\mathsf{P}},\mathbf{X}_{\mathsf{P}}) =&\|\mathbf{Y}-\bar{\mathbf{H}}\mathbf{X}\|^2\\ &+N_0\ln\det(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\mathbf{C})-N_0\mathbf{b}^{\mathsf{H}}(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\mathbf{C})^{-1}\mathbf{b}.\end{aligned}\tag{3.19}$$

Proof. See Appendix B.2. □

As already observed, the optimum receiver requires CDIR knowledge. Section 3.2.7 shows how CDIR can be obtained and the impact of its estimation on the receiver performance is addressed in Section 3.3.4.

3.2.4 Iterative computation of the decision metric

Following the approach of [36, 51], an iterative algorithm for the optimum receiver metric computation reducing substantially the computational effort is derived. For simplicity, the analysis focuses on the case of code word size matching the OFDM symbols size n_{F}, even though this approach can easily be extended to larger code word sizes.

Proposition 3.2.2. The metric expression (3.19) can be rewritten as:

$$\mu_{\mathsf{OR}}(\mathbf{X},\mathbf{Y},\mathbf{X}_p,\mathbf{Y}_p)=\sum_{k=1}^{n_{\mathsf{F}}}\Delta\mu_k(\mathbf{X}_k,\mathbf{Y}_k,\mathbf{X}_{\mathsf{P},k},\mathbf{Y}_{\mathsf{P},k}).\tag{3.20}$$

Here, the branch metric is defined as:

$$\begin{aligned}\Delta\mu_k(\mathbf{X}_k,\mathbf{Y}_k,\mathbf{X}_{\mathsf{P},k},\mathbf{Y}_{\mathsf{P},k}) \triangleq &\|\mathbf{Y}_k - \bar{\mathbf{H}}_k\mathbf{X}_k\|^2 + N_0 \ln\det(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \mathbf{U}_k\mathbf{\Lambda}_{k-1}\mathbf{U}_k^\mathsf{H}) \\ &- N_0\mathbf{b}_k^\mathsf{H}\mathbf{\Lambda}_k\mathbf{b}_k + N_0\mathbf{b}_{k-1}^\mathsf{H}\mathbf{\Lambda}_{k-1}\mathbf{b}_{k-1}, \end{aligned} \tag{3.21}$$

with

$$\begin{cases} \mathbf{U}_k & \triangleq & \mathbf{L}_k \otimes (\mathbf{F}^{1/2})_k \otimes \mathbf{R}^{1/2} \\ \mathbf{\Lambda}_k & \triangleq & \mathbf{\Lambda}_{k-1} + \Delta\mathbf{\Lambda} \\ \Delta\mathbf{\Lambda} & \triangleq & -\mathbf{\Lambda}_{k-1}\mathbf{U}_k^\mathsf{H}\left(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \mathbf{U}_k\mathbf{\Lambda}_{k-1}\mathbf{U}_k^\mathsf{H}\right)^{-1}\mathbf{U}_k\mathbf{\Lambda}_{k-1} \\ \mathbf{\Lambda_0} & \triangleq & \mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} \\ \mathbf{b}_k & \triangleq & \mathbf{b}_{k-1} + \Delta\mathbf{b} \\ \Delta\mathbf{b} & \triangleq & \mathrm{vec}\left\{(\mathbf{F}^{1/2})_k^\mathsf{H} \otimes \mathbf{R}^{1/2}\mathbf{B}_k\right\} \\ \mathbf{b_0} & \triangleq & \mathbf{0}_{n_\mathsf{R} n_\mathsf{T} n_\mathsf{F}}. \end{cases}$$

Matrix $\mathbf{L}_k$ is obtained from the Cholesky decomposition of $\mathbf{A}_k^\mathsf{T}$, i.e., $\mathbf{A}_k^\mathsf{T} = \mathbf{L}_k^\mathsf{H}\mathbf{L}_k$. $\mathbf{\Lambda_0}$ and $\mathbf{b_0}$ represent the initial values of $\mathbf{\Lambda}_k$ and $\mathbf{b}_k$, respectively, in the iterative algorithm. $\mathbf{0}_{n_\mathsf{R} n_\mathsf{T} n_\mathsf{F}}$ denotes the $n_\mathsf{R} n_\mathsf{T} n_\mathsf{F} \times n_\mathsf{R} n_\mathsf{T} n_\mathsf{F}$ all zeros matrix.

Proof. See Appendix B.3. □

3.2.5 Complexity

This section compares the complexity of the mismatched and the optimum receivers. Both receivers are based on Viterbi decoding, so that complexity is essentially due to branch metric computation.

We recall from section 2.2.5 the following properties

1. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, both sized $a \times b$, their sum $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$ requires ab complex sums.

2. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, sized $a \times b$ and $b \times c$, respectively, their product $\mathbf{Z} = \mathbf{XY}$ requires acb complex products and $ac(b-1)$ complex sums.

3. Given two matrices $\mathbf{X}$ and $\mathbf{Y}$, sized $a \times b$ and $c \times d$, respectively, their Kronecker product $\mathbf{Z} = \mathbf{X} \otimes \mathbf{Y}$ requires $abcd$ complex products.

4. The solution of a set of a linear equations in a unknowns, i.e. the solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ of $\mathbf{b} = \mathbf{Ax}$, where $\mathbf{x}$ contains the a unknowns, $\mathbf{A}$, sized $a \times a$, is the matrix of the coefficients and the vector $\mathbf{b}$ contains the a right–hand side quantities, performed by LU decomposition of $\mathbf{A}$ followed by forward and backward substitutions, requires $\frac{a^3}{3}$ complex products and $\frac{a^3}{3}$ complex sums for the LU decomposition, while the forward and backward substitutions require a^2 complex products and a^2 complex sums [57].

5. Hence, the solution $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ of $\mathbf{B} = \mathbf{A}\mathbf{X}$, with $\mathbf{X}$ and $\mathbf{B}$ sized $a \times b$, requires $\frac{a^3}{3} + ba^2$ complex products and $\frac{a^3}{3} + ba^2$ complex sums.

6. Moreover, once obtained the LU decomposition of $\mathbf{A}$, the evaluation of the determinant $|\mathbf{A}|$ requires $a - 1$ complex products.

7. Given a square, hermitian and positive–semidefinite matrix $\mathbf{X}$, sized $a \times a$ its Cholesky decomposition $\mathbf{X} = \mathbf{Z}^{\mathsf{H}}\mathbf{Z}$ requires $\frac{1}{6}a^3$ complex products and $\frac{1}{6}a^3$ complex sums [57].

8. Given a matrix $\mathbf{X}$ sized $a \times b$, its square Frobenius norm $\|\mathbf{X}\|^2$ requires ab complex products and $(a-1)(b-1)$ complex sums.

9. A complex sum requires 2 real sums.

10. A complex product requires 4 real products and 2 real sums.

Under a block fading assumption, the mismatched receiver requires the estimation of CSIR (i.e., the channel matrix $\mathbf{H}$) once at the beginning of the frame, whereas the optimum receiver requires the estimation of CDIR (i.e., the parameters $\bar{\mathbf{H}}_k$, K_k, $\mathbf{R}$, and $\mathbf{T}$). Since in the model considered, the subcarrier covariance matrix $\mathbf{F}$ depends only on the subcarrier separation and on the channel coherence bandwidth, it is assumed that it is known exactly at the receiver side, and the impact of its estimation error is not considered.

With the mismatched receiver,the channel matrices (3.12) for all subcarriers have to be estimated. The resulting computational cost amounts essentially to a matrix multiplication, since $\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}(\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}})^{-1}$ can be calculated off-line.

Therefore, the dominant complexity is the branch complexity which amounts to $n_{\mathsf{T}}n_{\mathsf{R}}N$ complex products and $n_{\mathsf{T}}n_{\mathsf{R}}N$ complex sums per branch metric as n_{T},n_{R},$n_{\mathsf{F}} \to \infty$. The details can be found in Table 3.1.

Table 3.1. Computational complexity of mismatched receiver branch metric evaluation (3.14)

No.	**Operation**	**Complex products No.**	**Complex sums No.**
(1)	$\widehat{\mathbf{H}}_k\mathbf{X}_k$	$n_{\mathsf{T}}n_{\mathsf{R}}N$	$(n_{\mathsf{T}}-1)n_{\mathsf{R}}N$
(2)	$\mathbf{Y}_k - (1)$	-	$n_{\mathsf{R}}N$
(3)	$\|(2)\|^2$	$n_{\mathsf{R}}N$	$(n_{\mathsf{R}}-1)(N-1)$
Dominant complexity		$n_{\mathsf{T}}n_{\mathsf{R}}N$	$n_{\mathsf{T}}n_{\mathsf{R}}N$

On the other hand, the optimum receiver does not need the explicit estimation of the channel matrix. Nevertheless, the computation of each branch metric computation (3.21) is larger as it requires the evaluation of the quantities reported in

Table 3.2. Notice that the product $\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}$ can be evaluated off-line.

Table 3.2. Computational complexity of optimum receiver branch metric evaluation (3.21)

No.	Operation	Complex products No.	Complex sums No.
(1)	$\|\mathbf{Y}_k - \bar{\mathbf{H}}_k\mathbf{X}_k\|^2$	$(n_{\mathsf{T}}+1)n_{\mathsf{R}}N$	$(n_{\mathsf{T}}+1)n_{\mathsf{R}}N - n_{\mathsf{R}} - N + 1$
	$\mathbf{A}_k$		
(2)	$\mathbf{X}_k\mathbf{X}_k^{\mathsf{H}}$	$n_{\mathsf{T}}^2 N$	$n_{\mathsf{T}}^2(N-1)$
(3)	$(2) + \mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}$	$-$	n_{T}^2
(4)	$\mathbf{T}^{1/2}(3)$	n_{T}^3	$n_{\mathsf{T}}^2(n_{\mathsf{T}}-1)$
(5)	$(4)\mathbf{T}^{1/2}$	n_{T}^3	$n_{\mathsf{T}}^2(n_{\mathsf{T}}-1)$
(6)	$\mathbf{L}_k$	$\frac{1}{6}n_{\mathsf{T}}^3$	$\frac{1}{6}n_{\mathsf{T}}^3$
(7)	$\mathbf{U}_k$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}$	$-$
	$\Delta\mathbf{\Lambda}$		
(8)	$\mathbf{U}_k$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}$	$-$
(9)	$\mathbf{\Lambda}_-(8)^{\mathsf{H}}$	$n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}^2$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}(n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}-1)$
(10)	$\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}} + \mathbf{U}_k(9)$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}$
(11)	$(10)^{-1}(9)^{\mathsf{H}}$	$\frac{n_{\mathsf{T}}^3 n_{\mathsf{R}}^3}{3} + n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}$	$\frac{n_{\mathsf{T}}^3 n_{\mathsf{R}}^3}{3} + n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}$
(12)	$(9)(11)$	$n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}^2$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}^2(n_{\mathsf{T}}n_{\mathsf{R}}-1)$
(13)	$\mathbf{\Lambda} = \mathbf{\Lambda}_- + (12)$	$-$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}^2$
	$\mathbf{B}_k$		
(14)	$\mathbf{Y}_k\mathbf{X}_k^{\mathsf{H}}$	$n_{\mathsf{T}}n_{\mathsf{R}}N$	$n_{\mathsf{T}}n_{\mathsf{R}}(N-1)$
(15)	$\bar{\mathbf{H}}_k(2)$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}$	$n_{\mathsf{T}}(n_{\mathsf{T}}-1)n_{\mathsf{R}}$
(16)	$\mathbf{Y}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}$	$n_{\mathsf{T}}n_{\mathsf{R}}P$	$n_{\mathsf{T}}n_{\mathsf{R}}(P-1)$
(17)	$\bar{\mathbf{H}}_k\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}}$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}$	$n_{\mathsf{T}}(n_{\mathsf{T}}-1)n_{\mathsf{R}}$
(18)	$(14)-(15)+(16)-(17)$	$-$	$3n_{\mathsf{T}}n_{\mathsf{R}}$
(19)	$(18)\mathbf{T}^{1/2}$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}$	$n_{\mathsf{T}}(n_{\mathsf{T}}-1)n_{\mathsf{R}}$
(20)	$\Delta\mathbf{b}$	$n_{\mathsf{T}}n_{\mathsf{R}}^2 n_{\mathsf{F}}$	$n_{\mathsf{T}}n_{\mathsf{R}}(n_{\mathsf{R}}-1)n_{\mathsf{F}}$
(21)	$\mathbf{b} = \mathbf{b}_- + (20)$	$-$	$n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}$
(22)	$(13)(21)$	$n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}^2$	$n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}(n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}-1)$
(23)	$(21)^{\mathsf{H}}(22)$	$n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}$	$n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}-1$
(24)	$\det\{(10)\}$	$n_{\mathsf{T}}n_{\mathsf{R}}-1$	$-$
Dominant complexity		$2n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}^2$	$2n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}^2$

Summarizing, the overall decoding complexity amounts to $2n_{\mathsf{T}}^3 n_{\mathsf{R}}^3 n_{\mathsf{F}}^2$ operations per branch metric as n_{T},n_{R},$n_{\mathsf{F}} \to \infty$.

Thus, the optimum receiver complexity increase with respect to the mismatched receiver can be approximated by $2N^{-1}n_{\mathsf{T}}^2 n_{\mathsf{R}}^2 n_{\mathsf{F}}^2$.

3.2.6 Spectral Approximation of the decision metric

It has been shown in section 2.2.5 that the complexity of the optimum metric evaluation depends on n_{F}^2. Since the number of subcarriers n_{F} can be large (it is assumed in Section 3.3 that n_{F}=117), complexity can be an important issue for practical implementation of the proposed receiver. However, a primary simplification derives by using the iterative metric (3.20) instead of (3.19), which reduces the order of the matrix inversion from $n_{\mathsf{T}} n_{\mathsf{R}} n_{\mathsf{F}}$ to $n_{\mathsf{R}} n_{\mathsf{T}}$.

A further simplification can be obtained when the ordered sequence of eigenvalues of the matrix $\mathbf{F}$ decays sufficiently fast. This opportunity is precisely addressed by Proposition 3.2.3 and Corollary 3.2.1.

Proposition 3.2.3. Let

$$\mathbf{F} = \mathbf{V}\mathbf{D}^2\mathbf{V}^{\mathsf{H}}$$

be the sorted orthogonal decomposition of $\mathbf{F}$, such that the diagonal entries of $\mathbf{D}$ are a nonincreasing sequence. Let m_{F} be an integer number with $0 < m_{\mathsf{F}} < n_{\mathsf{F}}$.

Define the column vectors

$$\mathbf{g}_k^{m_{\mathsf{F}}} \triangleq (\mathbf{V}\mathbf{D}_{m_{\mathsf{F}}})_k^{\mathsf{H}}$$

and

$$\overline{\mathbf{g}}_k^{m_{\mathsf{F}}} \triangleq (\mathbf{V}(\mathbf{D} - \mathbf{D}_{m_{\mathsf{F}}}))_k^{\mathsf{H}}$$

with

$$\mathbf{D}_{m_{\mathsf{F}}} \triangleq \mathbf{D} \cdot (\mathbf{I}_{m_{\mathsf{F}}} \oplus \mathbf{0}_{n_{\mathsf{F}} - m_{\mathsf{F}}})$$

being the diagonal matrix containing the largest m_{F} eigenvalues f_i of $\mathbf{F}$.

Now set

$$\begin{cases} \mathbf{b}_{m_{\mathsf{F}}} & \triangleq & \displaystyle\sum_{k=1}^{n_{\mathsf{F}}} \mathrm{vec}\left\{\mathbf{g}_k^{m_{\mathsf{F}}} \otimes \mathbf{R}^{1/2}\mathbf{B}_k\right\} \\ \mathbf{C}_{m_{\mathsf{F}}} & \triangleq & \displaystyle\sum_{k=1}^{n_{\mathsf{F}}} \mathbf{A}_k^{\mathsf{T}} \otimes (\mathbf{g}_k^{m_{\mathsf{F}}} \mathbf{g}_k^{m_{\mathsf{F}}\mathsf{H}}) \otimes \mathbf{R} \\ \overline{\mathbf{b}}_{m_{\mathsf{F}}} & \triangleq & \mathbf{b} - \mathbf{b}_{m_{\mathsf{F}}} \\ \overline{\mathbf{C}}_{m_{\mathsf{F}}} & \triangleq & \mathbf{C} - \mathbf{C}_{m_{\mathsf{F}}} \end{cases}$$

and define the metric components:

$$\begin{array}{lcllcl} \mu_{\mathsf{OR}}^{(1)} & \triangleq & N_0 \ln\det(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}} n_{\mathsf{F}}} + \mathbf{C}) & \mu_{\mathsf{OR}}^{(2)} & \triangleq & N_0 \mathbf{b}^{\mathsf{H}}(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}} n_{\mathsf{F}}} + \mathbf{C})^{-1}\mathbf{b} \\ \widetilde{\mu}_{\mathsf{OR}}^{(1)} & \triangleq & N_0 \ln\det(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}} n_{\mathsf{F}}} + \mathbf{C}_{m_{\mathsf{F}}}) & \widetilde{\mu}_{\mathsf{OR}}^{(2)} & \triangleq & N_0 \mathbf{b}_{m_{\mathsf{F}}}^{\mathsf{H}}(\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}} n_{\mathsf{F}}} + \mathbf{C}_{m_{\mathsf{F}}})^{-1}\mathbf{b}_{m_{\mathsf{F}}} \ . \end{array}$$

Then μ_{OR} can be approximated by

$$\begin{aligned} \mu_{\mathsf{OR}} &= \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 + \mu_{\mathsf{OR}}^{(1)} - \mu_{\mathsf{OR}}^{(2)} \\ \approx &\widetilde{\mu}_{\mathsf{OR}} \triangleq \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 + \widetilde{\mu}_{\mathsf{OR}}^{(1)} - \widetilde{\mu}_{\mathsf{OR}}^{(2)} \ . \end{aligned}$$

If

$$|||\overline{\mathbf{C}}_{m_\mathsf{F}}||| \cdot |||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})^{-1}||| < 1$$

the approximation errors are bounded by:

$$\begin{aligned}
\frac{|\mu_{\mathsf{OR}}^{(1)} - \widetilde{\mu}_{\mathsf{OR}}^{(1)}|}{N_0} \leq & n_\mathsf{F} \left| \ln\left(1 - \kappa \frac{|||\overline{\mathbf{C}}_{m_\mathsf{F}}|||}{|||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})|||}\right)\right| \\
\frac{|\mu_{\mathsf{OR}}^{(2)} - \widetilde{\mu}_{\mathsf{OR}}^{(2)}|}{N_0} \leq & \mu_{\mathsf{OR}}^{(2)} \frac{\kappa}{1 - \kappa |||\overline{\mathbf{C}}_{m_\mathsf{F}}|||/|||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})|||} \cdot \frac{|||\overline{\mathbf{C}}_{m_\mathsf{F}}|||}{|||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})|||} + \|\overline{\mathbf{b}}_{m_\mathsf{F}}\|^2
\end{aligned} \tag{3.22}$$

with

$$\kappa \triangleq |||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})||| \cdot |||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})^{-1}|||$$

being the *condition number* of $(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})$. Moreover, $\|\overline{\mathbf{b}}_{m_\mathsf{F}}\|$ and $|||\overline{\mathbf{C}}_{m_\mathsf{F}}|||$ are bounded by:

$$\begin{aligned}
\|\overline{\mathbf{b}}_{m_\mathsf{F}}\|^2 \leq & \sqrt{n_\mathsf{F}} f_{m_\mathsf{F}+1} \sqrt{\sum_{k,l=1}^{n_\mathsf{F}} |\mathrm{Tr}\left(\mathbf{B}_k^\mathsf{H} \mathbf{R} \mathbf{B}_l\right)|^2} \\
|||\overline{\mathbf{C}}_{m_\mathsf{F}}||| \leq & \left(f_{m_\mathsf{F}+1} + 2\sqrt{|||\mathbf{F}||| \cdot f_{m_\mathsf{F}+1}}\right) \cdot |||\mathbf{R}||| \cdot \sum_{k=1}^{n_\mathsf{F}} |||\mathbf{A}_k|||\ .
\end{aligned} \tag{3.23}$$

Proof. see Appendix B.4. □

Corollary 3.2.1. Removing the redundant dimensions of $\mathbf{D}_{m_\mathsf{F}}$, which do not contribute to the metric computation, by setting

$$\widetilde{\mathbf{D}}_{m_\mathsf{F}} \triangleq \mathbf{D}_{m_\mathsf{F}}[\mathbf{I}_{m_\mathsf{F}}, \mathbf{0}_{m_\mathsf{F} \times (n_\mathsf{F} - m_\mathsf{F})}]^\mathsf{T}$$

gives:

$$\widetilde{\mu}_{\mathsf{OR}}^{(1)} = N_0 \ln\det(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} m_\mathsf{F}} + \widetilde{\mathbf{C}}_{m_\mathsf{F}}) \qquad \widetilde{\mu}_{\mathsf{OR}}^{(2)} = N_0 \widetilde{\mathbf{b}}_{m_\mathsf{F}}^\mathsf{H} (\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} m_\mathsf{F}} + \widetilde{\mathbf{C}}_{m_\mathsf{F}})^{-1} \widetilde{\mathbf{b}}_{m_\mathsf{F}}.$$

with

$$\begin{cases}
\widetilde{\mathbf{b}}_{m_\mathsf{F}} &= \displaystyle\sum_{k=1}^{n_\mathsf{F}} \mathrm{vec}\left\{(\mathbf{V}\mathbf{D}_{m_\mathsf{F}})_k^\mathsf{H} \otimes \mathbf{R}^{1/2} \mathbf{B}_k\right\} \\
\widetilde{\mathbf{C}}_{m_\mathsf{F}} &= \displaystyle\sum_{k=1}^{n_\mathsf{F}} \mathbf{A}_k^\mathsf{T} \otimes ((\mathbf{V}\widetilde{\mathbf{D}}_{m_\mathsf{F}})_k^\mathsf{H} (\mathbf{V}\widetilde{\mathbf{D}}_{m_\mathsf{F}})_k) \otimes \mathbf{R}.
\end{cases}$$

Note that the validity of the bounds in equations (3.22) is based on the assumption that

$$|||\overline{\mathbf{C}}_{m_\mathsf{F}}||| \cdot |||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})^{-1}||| < 1$$

. Since (plainly)

$$|||(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{C})^{-1}||| < 1$$

it is sufficient to assume that $|||\overline{\mathbf{C}}_{m_\mathsf{F}}||| < 1$. Solving the quadratic equation for the second of equations (3.23) gives in turn a sufficient condition for the validity of the bounds in equations (3.22) depending of $\sqrt{f_{m_\mathsf{F}+1}}$.

Summarizing, Corollary 3.2.1 states that the dimension of the matrix $\mathbf{C}$ and the vector $\mathbf{b}$ in the approximative metric computation can be reduced to $n_\mathsf{T} n_\mathsf{R} m_\mathsf{F} \times n_\mathsf{T} n_\mathsf{R} m_\mathsf{F}$ and $n_\mathsf{T} n_\mathsf{R} m_\mathsf{F}$, respectively. Proposition 3.2.3 guarantees that the error in the metric approximation is of order $\sqrt{f_{m_\mathsf{F}+1}}$, provided that $\sqrt{f_{m_\mathsf{F}+1}}$ is small enough.

Thus, if $m_\mathsf{F} \ll n_\mathsf{F}$ this leads to a considerable computational gain. In the numerical simulations, Section 3.3, the correlation between carriers will be modeled by a matrix with Toeplitz structure to be defined in equation (3.25). As shown in Figure 3.3 most of the energy of $\mathbf{F}^{1/2}$ is concentrated in the first few eigenvalues.

3.2.7 CDIR estimation

CDIR knowledge for the channel model considered consists of the following parameters:

1. the average channel matrices $\bar{\mathbf{H}}_k$ (for $k = 1, \ldots, n_\mathsf{F}$);
2. the Rice factor K_k (for $k = 1, \ldots, n_\mathsf{F}$);
3. the covariance matrices $\mathbf{R}$ and $\mathbf{T}$.

As already mentioned, $\mathbf{F}$ is assumed to be known at the receiver.

A simple estimation algorithm for the channel parameters is given as follows. During the ith fading block, the following channel equation corresponding to the kth subcarrier holds:

$$\mathbf{Y}_{\mathsf{P},k}^{(i)} = \mathbf{H}_k^{(i)} \mathbf{X}_{\mathsf{P},k} + \mathbf{Z}_{\mathsf{P},k}^{(i)} \qquad k = 1, \ldots, n_\mathsf{F}, \;\; i = 1,2, \ldots$$

As in Section 3.2.2, the ML estimate of $\mathbf{H}_k^{(i)}$ is given by

$$\widehat{\mathbf{H}}_k^{(i)} = \mathbf{Y}_{\mathsf{P},k}^{(i)} \mathbf{X}_{\mathsf{P},k}^{\mathsf{H}} (\mathbf{X}_{\mathsf{P},k} \mathbf{X}_{\mathsf{P},k}^{\mathsf{H}})^{-1} \; .$$

Since the joint ML estimation is very complex, it is preferable to resort to a simpler approach based on the knowledge of the L estimates $\widehat{\mathbf{H}}_k^{(1)}, \ldots, \widehat{\mathbf{H}}_k^{(L)}$.

- The average channel matrices are estimated as

$$\widehat{\mathbf{E}}_k = \frac{1}{L}\sum_{i=1}^{L}\widehat{\mathbf{H}}_k^{(i)} \qquad\qquad k = 1,\ldots,n_{\mathsf{F}}.$$

- The Rice factors K_k are estimated by

$$\widehat{K}_k = \frac{\|\widehat{\mathbf{E}}_k\|^2}{\frac{1}{L}\sum_{i=1}^{L}\|\widehat{\mathbf{H}}_k^{(i)} - \widehat{\mathbf{E}}_k\|^2} \qquad\qquad k = 1,\ldots,n_{\mathsf{F}}.$$

- The estimates of $\mathbf{R}$ and $\mathbf{T}$ derive from the following equations obtained from (3.7):

$$\begin{aligned}
\sum_{k=1}^{n_{\mathsf{F}}}\mathbb{E}\left[\mathbf{H}_k\mathbf{H}_k^{\mathsf{H}}\right] &= \sum_{k=1}^{n_{\mathsf{F}}}\bar{\mathbf{H}}_k\bar{\mathbf{H}}_k^{\mathsf{H}} + \mathrm{Tr}\left(\mathbf{F}\right)\mathrm{Tr}\left(\mathbf{T}\right)\mathbf{R}\\
\sum_{k=1}^{n_{\mathsf{F}}}\mathbb{E}\left[\mathbf{H}_k^{\mathsf{H}}\mathbf{H}_k\right] &= \sum_{k=1}^{n_{\mathsf{F}}}\bar{\mathbf{H}}_k^{\mathsf{H}}\bar{\mathbf{H}}_k + \mathrm{Tr}\left(\mathbf{F}\right)\mathrm{Tr}\left(\mathbf{R}\right)\mathbf{T}\\
\sum_{k=1}^{n_{\mathsf{F}}}\mathbb{E}\left[\|\mathbf{H}_k\|^2\right] &= \sum_{k=1}^{n_{\mathsf{F}}}\|\bar{\mathbf{H}}_k\|^2 + \mathrm{Tr}\left(\mathbf{F}\right)\mathrm{Tr}\left(\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{T}\right)
\end{aligned}$$

  From these equations the following estimates are derived:

$$\begin{aligned}
\widehat{\mathbf{R}} &\triangleq \frac{1}{L\hat{f}\sqrt{\hat{q}}}\sum_{i=1}^{L}\sum_{k=1}^{n_{\mathsf{F}}}\left[\widehat{\mathbf{H}}_k^{(i)}\left(\widehat{\mathbf{H}}_k^{(i)}\right)^{\mathsf{H}} - \widehat{\mathbf{E}}_k\widehat{\mathbf{E}}_k^{\mathsf{H}}\right]\\
\widehat{\mathbf{T}} &\triangleq \frac{1}{L\hat{f}\sqrt{\hat{q}}}\sum_{i=1}^{L}\sum_{k=1}^{n_{\mathsf{F}}}\left[\left(\widehat{\mathbf{H}}_k^{(i)}\right)^{\mathsf{H}}\widehat{\mathbf{H}}_k^{(i)} - \widehat{\mathbf{E}}_k^{\mathsf{H}}\widehat{\mathbf{E}}_k\right]
\end{aligned}$$

  where $\hat{f} \triangleq \mathrm{Tr}\left(\mathbf{F}\right)$ and

$$\hat{q} \triangleq \frac{1}{L\hat{f}}\sum_{i=1}^{L}\sum_{k=1}^{n_{\mathsf{F}}}\left[\|\widehat{\mathbf{H}}_k^{(i)}\|^2 - \|\widehat{\mathbf{E}}_k\|^2\right].$$

  Here, it was assumed implicitly that the traces of $\mathbf{R}$ and $\mathbf{T}$ are both equal to the square root of the estimate $\hat{q}$ of their product.

  It is plain to see that the two matrix estimates $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{T}}$ are positive semidefinite.

The receiver will then operate according to two different modes:

- **Training mode.** The first L blocks are used to estimate the channel parameters in parallel to the normal operation of the mismatched receiver.
- **Operating mode.** The next blocks are processed by using the optimum receiver with the parameters estimated during all the fading blocks preceding the current one.

Remark 3.2.2. Our definitions of training and operating modes apply to full channel frames and not to the partial frames assigned to pilots or data.

It is worth to state that information symbols are transmitted during both modes since the CDIR parameter estimation algorithm is assumed to run in parallel with the decoding algorithm. Therefore, there is no throughput loss because of parameter acquisition, even though the error performance is expected to be worse during the training mode (when the mismatched receiver is used) than during the operating mode (when the optimum receiver comes into play, based on the estimated CDIR).

Since CDIR estimation is performed continuously and not only during the training mode, the receiver is insensitive to its duration. This allows to ignore the problem of assessing the minimum length of the training allowing for reliable CDIR estimation.

Remark 3.2.3. Since the number of parameters to estimate grows approximately as $n_{\mathsf{R}}^2 + n_{\mathsf{T}}^2$, the duration of the training mode to grows approximately as $\max\{n_{\mathsf{T}}, n_{\mathsf{R}}\}$. However, in our numerical examples, this effect is unnoticed a 2×2 MIMO system is considered.

3.3 Numerical Results

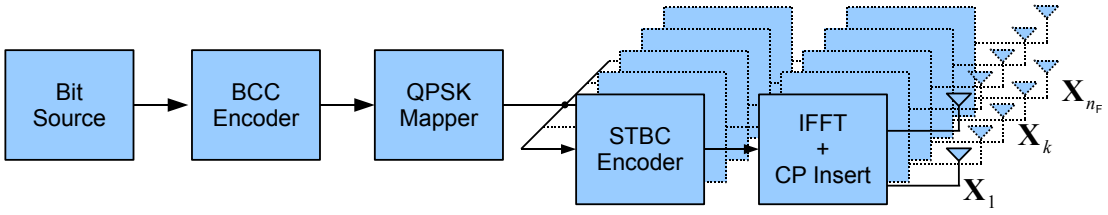


Figure 3.2. The MIMO-OFDM transmitting scheme

In this section the MIMO-OFDM system with $n_{\mathsf{T}} = 2$ transmit and $n_{\mathsf{R}} = 2$ receive antennas depicted in Figure 3.2 is considered.

It is assumed that the transmitter complies with the 802.11n draft 2.0 standard [49]. More specifically, the Modulation and Coding Scheme (MCS) 1 with 40 MHz channel (Greenfield format) and the (optional) Space-Time Block Code STBC [4] is chosen. The number of subcarriers is $n_{\mathsf{F}} = 117$ (108 dedicated to data transmission[1]). The random bit stream is fed to a Binary Convolutional Encoder with standard generator polynomials

$$g_0 = (133)_8$$
$$g_1 = (171)_8$$

of rate $R = 1/2$. Coded bits are modulated with QPSK and Gray mapping. Complex QPSK symbols are pairwise mapped on the data subcarriers accounting for the presence of the STBC. For every subcarrier the standard STBC implements the following mapping from symbol pairs $(x_{k,1}\ x_{k,2})$ to 2×2 matrices:

$$(x_{k,1}\ x_{k,2}) \mapsto \begin{pmatrix} x_{k,1} & x_{k,2} \\ -x_{k,2}^* & x_{k,1}^* \end{pmatrix} \tag{3.24}$$

It is assumed that the maximum number of bits (208 data, 6 tail, and 2 padding bits) is transmitted for each OFDM symbol. All subcarrier matrices $\mathbf{X}_k$ have length $N = 2$. Every data symbol matrix is preceded by an orthogonal pilot symbol matrix, so that

$$\mathbf{X}_{\mathsf{P},k}\mathbf{X}_{\mathsf{P},k}^{\mathsf{H}} \propto \mathbf{I}_{n_{\mathsf{T}}}$$

for $k = 1, \ldots, n_{\mathsf{F}}$.

[1] See [49] for the details

According to the standard specifications for $n_{\mathsf{T}} = 2$ transmit antennas, set $P = 2$. Data and pilot matrices are in order to have unit energy over one symbol period.

Simulations have been carried out for the mismatched and optimum receivers. The genie receiver has been considered as a benchmark.

Rayleigh and Rician fading channels are considered, correlated among subcarriers, with and without spatial correlation at the transmitter and the receiver. The Rician factor K_k are assumed to be all equal to K.

Transmit and receive correlation matrices $\mathbf{T}$ and $\mathbf{R}$ are assumed to follow the exponential model proposed by [11]:

$$(\mathbf{T})_{ij} = \begin{cases} \rho_{\mathsf{T}}^{j-i}, & i \leq j \\ (\mathbf{T})_{ji}^{*}, & i > j \end{cases}$$

and

$$(\mathbf{R})_{ij} = \begin{cases} \rho_{\mathsf{R}}^{j-i}, & i \leq j \\ (\mathbf{R})_{ji}^{*}, & i > j \end{cases}$$

where ρ_{T} and ρ_{R} represent complex spatial correlation coefficients of neighboring antennas.

Subcarrier correlation is modeled according to [63] as follows:

$$(\mathbf{F})_{ij} = \left[\sqrt{1 + \left(|i-j| \frac{\Delta f}{\Delta f_c} \right)^2} \right]^{-1} \tag{3.25}$$

where $\Delta f = 312.5$ kHz is the separation between carriers set by the standard and $\Delta f_c = 10$ MHz is the coherence bandwidth of the channel.

For the sake of simplicity and following the assumption in *e.g.* [59, 17, 13], the average channel matrices are assumed to have constant entries, i.e.,

$$\bar{\mathbf{H}}_k = \sqrt{\frac{K}{K+1}} \cdot \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

3.3.1 Numerical validation of spectral approximation

The numerical validity of the spectral approximation discussed in Section 3.2.6 is illustrated in Figure 3.3, which plots the square Frobenius norms of $\mathbf{F}^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathsf{H}}$ and $\widetilde{\mathbf{F}}^{1/2} = \mathbf{V}\mathbf{D}_{m_{\mathsf{F}}}\mathbf{V}^{\mathsf{H}}$ versus the number of largest eigenvalues m_{F} considered.

It can be noticed that the first 8 largest eigenvalues encompass approximately the whole energy of $\mathbf{F}^{1/2}$. Thus, the effect of the spectral approximation in the following three cases

1. below the threshold ($m_{\mathsf{F}} = 4$);
2. at the threshold ($m_{\mathsf{F}} = 8$);
3. above the threshold ($m_{\mathsf{F}} = 16$)

was studied.

Performance results are reported in terms of FER (Frame Error Rate) versus SNR in dB with:

$$\mathsf{SNR} = n_{\mathsf{T}} \frac{E_{\mathsf{S}}}{N_0}, \tag{3.26}$$

under the assumption of constant Rician factor K and SNR among subcarriers.

3.3.2 Optimum receiver with Rayleigh fading: impact of m_{F}

Fig. 3.4 plots the error performance in the uncorrelated Rayleigh case ($\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.0$ and $K = 0$) versus the SNR. The diagrams show results for the genie, mismatched, and optimum receivers with a number of eigenvalues considered at the receiver $m_{\mathsf{F}} = 4,8,16$.

It can be noticed that the error performance corresponding to $m_{\mathsf{F}} = 4$ loses up to 1 dB from the case of $m_{\mathsf{F}} = 8$, which practically overlaps with the case of $m_{\mathsf{F}} = 16$. The curves also show a diversity loss corresponding to $m_{\mathsf{F}} = 4$. As a result, it turns out that there is no point in this case for considering more than the first $m_{\mathsf{F}} = 8$ largest eigenvalues in the optimum receiver as the extra effort would not yield a significant return.

Moreover, it can be noticed that the optimum receiver has a penalty of about 0.7 dB with respect to the genie receiver but it gains more than 3 dB with respect to the standard mismatched receiver.

Similar results are reported in Figure 3.5 for the spatially-correlated ($\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.7$) Rayleigh case. It can be noticed that this level of spatial correlation introduces a performance degradation of about 2.5 dB with respect to the uncorrelated case.

Also in this case, setting $m_{\mathsf{F}} = 8$ yields the best performance trade-off. The optimum receiver has a loss of about 0.7 dB from the genie receiver and a gain of 3.5 dB with respect to the mismatched receiver.

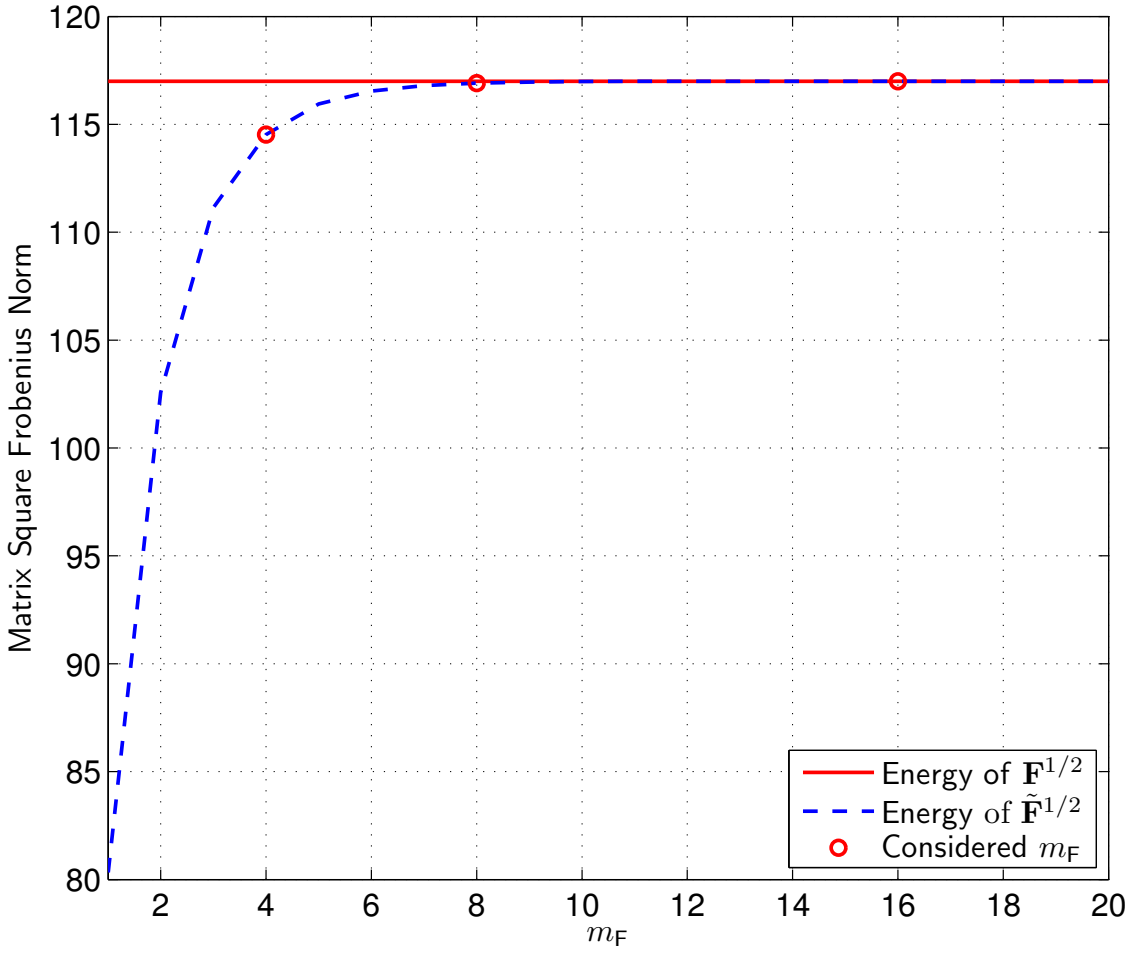


Figure 3.3. Energy of $\mathbf{F}^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^{\mathsf{H}}$ and $\widetilde{\mathbf{F}}^{1/2} \triangleq \mathbf{V}\mathbf{D}_{m_{\mathsf{F}}}\mathbf{V}^{\mathsf{H}}$ vs. m_{F}. The considered values of $m_{\mathsf{F}} = 4{,}8{,}16$ are highlighted.

Therefore, our numerical results presented in Figs. 3.4 and 3.5 are consistent with the analysis carried out in Section 3.2.6. In fact, a substantial complexity reduction can be obtained by considering $\widetilde{\mathbf{F}}^{1/2}$ rather than $\mathbf{F}^{1/2}$ with negligible penalty on the error performance provided that the number of largest eigenvalues considered is sufficiently large (in the case considered, $m_{\mathsf{F}} = 8$).

The impact of m_{F} on receiver complexity is also illustrated by the results in Table 3.3, which report the CPU time per decoded frame over an Intel® Xeon™3.20 GHz machine with the three types of receivers considered in this work and different values of m_{F} for the optimum receiver.

It can be noticed that, without spectral approximation, the time complexity would exceed 100 times the complexity of the mismatched receiver, whereas the recommended choice of $m_{\mathsf{F}} = 8$ implies an overhead of 60%, which is traded off for a 3 dB SNR gain at fixed error performance.

Similar results hold in the Rician case so that the value $m_{\mathsf{F}} = 8$ will be fixed in the following sections.

Table 3.3. Time complexity of different receiver schemes

Receiver		**Time**
Mismatched		1
Optimum	$m_{\mathsf{F}} = 4$	1.27
Optimum	$m_{\mathsf{F}} = 8$	1.60
Optimum	$m_{\mathsf{F}} = 16$	3.38
Optimum	$m_{\mathsf{F}} = n_{\mathsf{F}}$ (no approximation)	153

3.3.3 Optimum receiver with Riciang fading: Effects of correlation and the Rice factor

In this section a MIMO-OFDM system with Rician factors $K = 0$ dB and $K = 10$ dB are considered.

Figure 3.6 shows the error performance with $K = 0$ dB and $\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.0$. The curves show that the optimum receiver has a penalty of about 0.5 dB with respect to the genie receiver and a gain of about 3.5 dB with respect to the Mismatched receiver.

Spatial correlation corresponding to $\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.7$ implies an overall penalty of about 4.5 dB with respect to the uncorrelated case as is shown in Figure 3.7. In this case, the optimum receiver loses about 0.7 dB from the genie receiver and gains about 3.5 dB with respect to the Mismatched receiver.

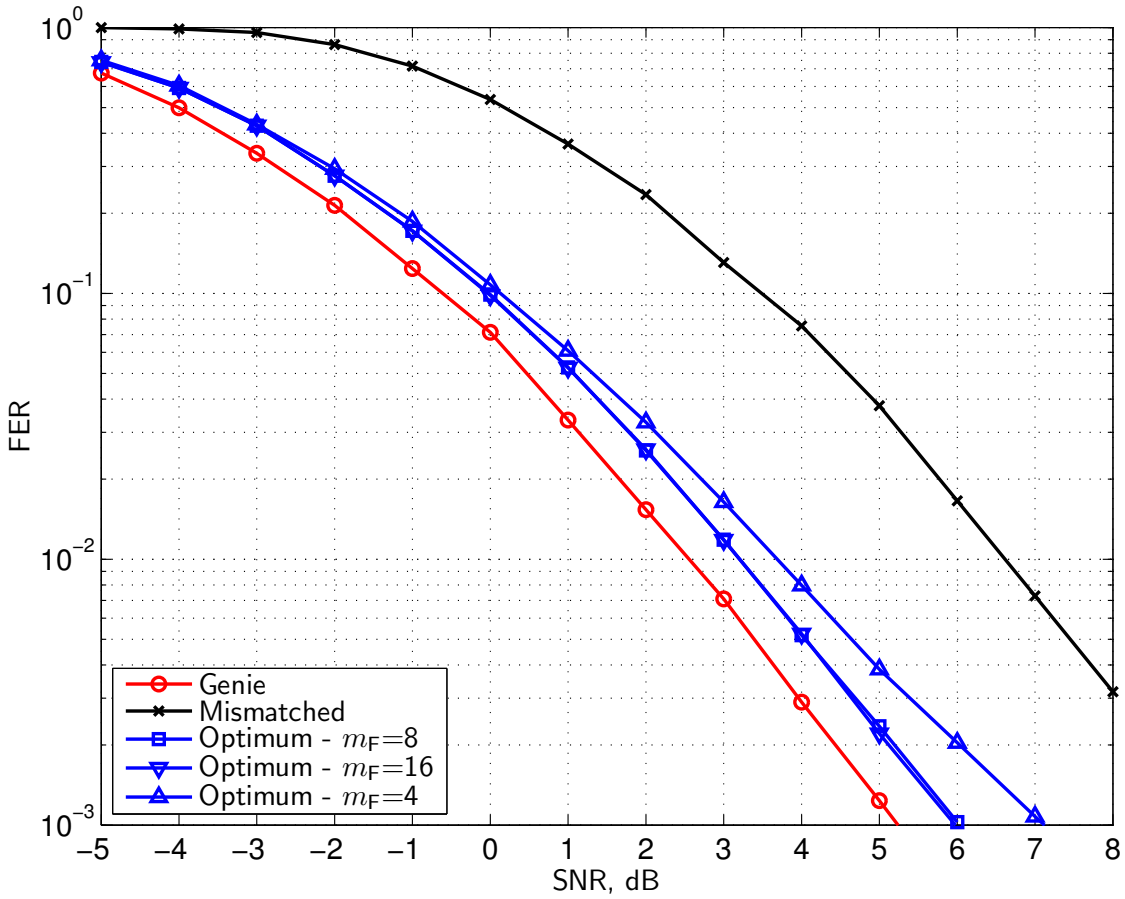


Figure 3.4. FER vs. SNR of the MIMO-OFDM system with $K = 0$ (Rayleigh fading), $\rho_R = \rho_T = 0.0$. Genie, mismatched and optimum ($m_F = 4,8,16$) receivers.

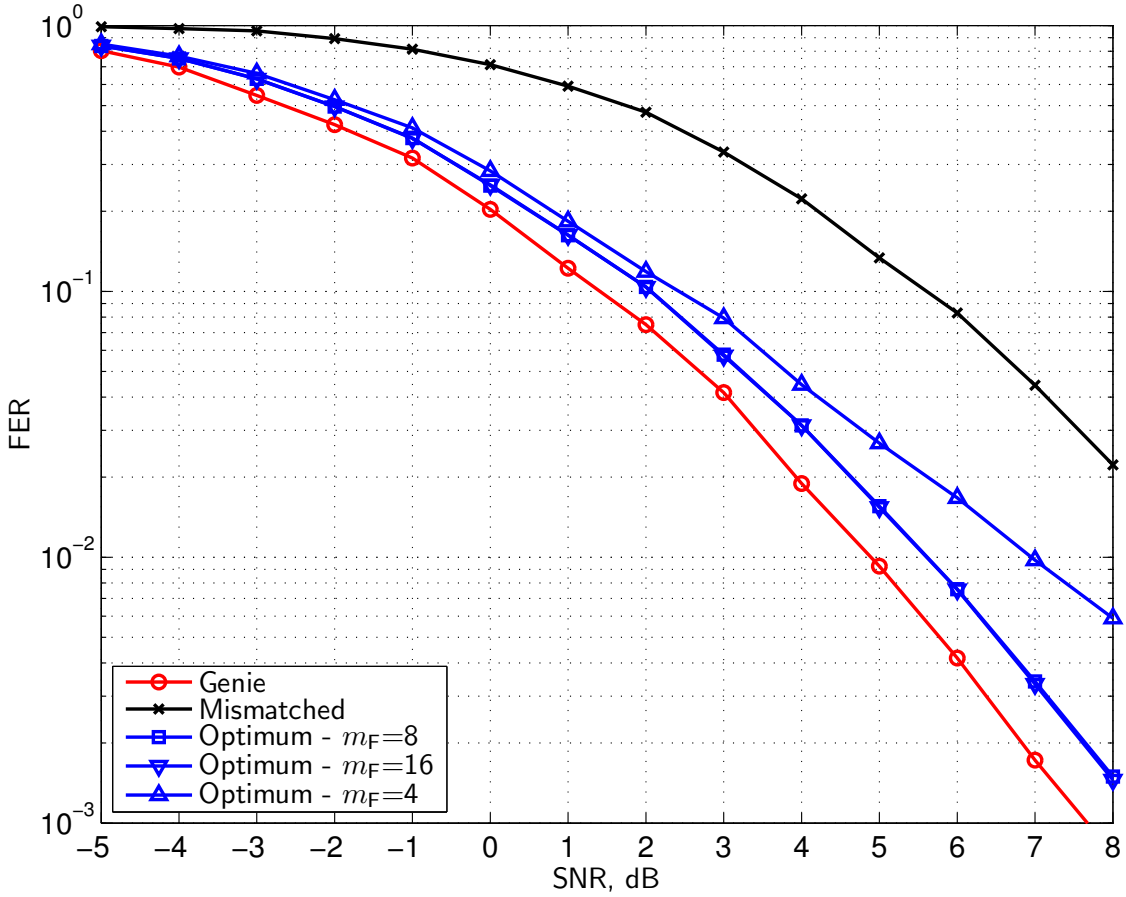


Figure 3.5. FER vs. SNR of the MIMO-OFDM system with $K = 0$ (Rayleigh fading), $\rho_R = \rho_T = 0.7$. Genie, mismatched and optimum ($m_F = 4,8,16$) receivers.

Similar effects are noticed when the Rice factor increases to $K = 10$ dB. Figure 3.8 shows the error performance in the uncorrelated case, where the optimum receiver has a slim penalty of 0.2 dB with respect to the genie receiver, while it gains about 3.5 dB with respect to the Mismatched receiver.

The spatially correlated case with $\rho_{\mathsf{R}} = \rho_{\mathsf{T}} = 0.7$ is accounted for in Figure 3.9. Here, it can be noticed that the optimum receiver loses about 0.2 dB from the genie receiver and gains almost 4 dB with respect to the Mismatched receiver.

Numerical results considered so far were based on the assumption of perfect CDIR, that is, the receiver knows exactly the statistic parameters of the channel model. The effect of estimating these channel parameters is addressed by the following section.

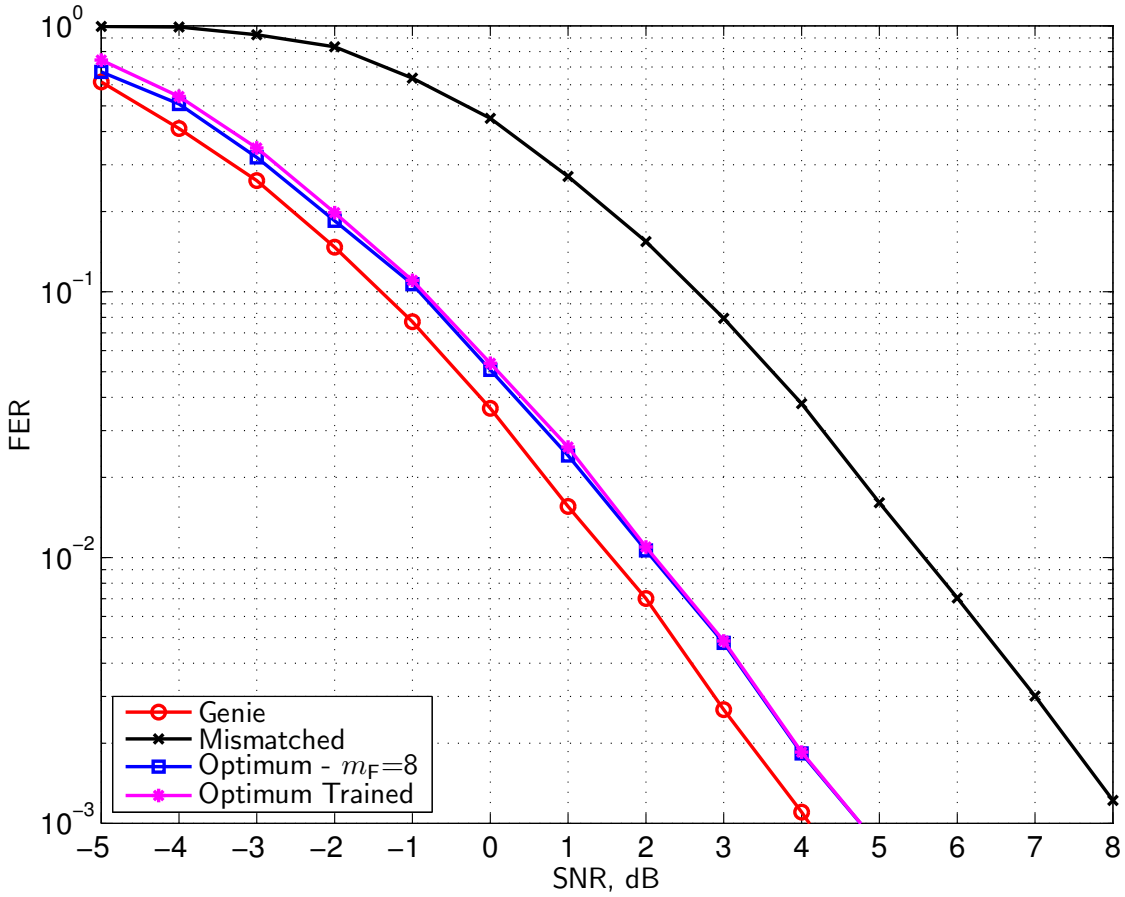


Figure 3.6. FER vs. SNR of the MIMO-OFDM system with $K = 0$ dB (Rice fading), $\rho_R = \rho_T = 0.0$. Genie, mismatched, optimum and optimum trained receivers.

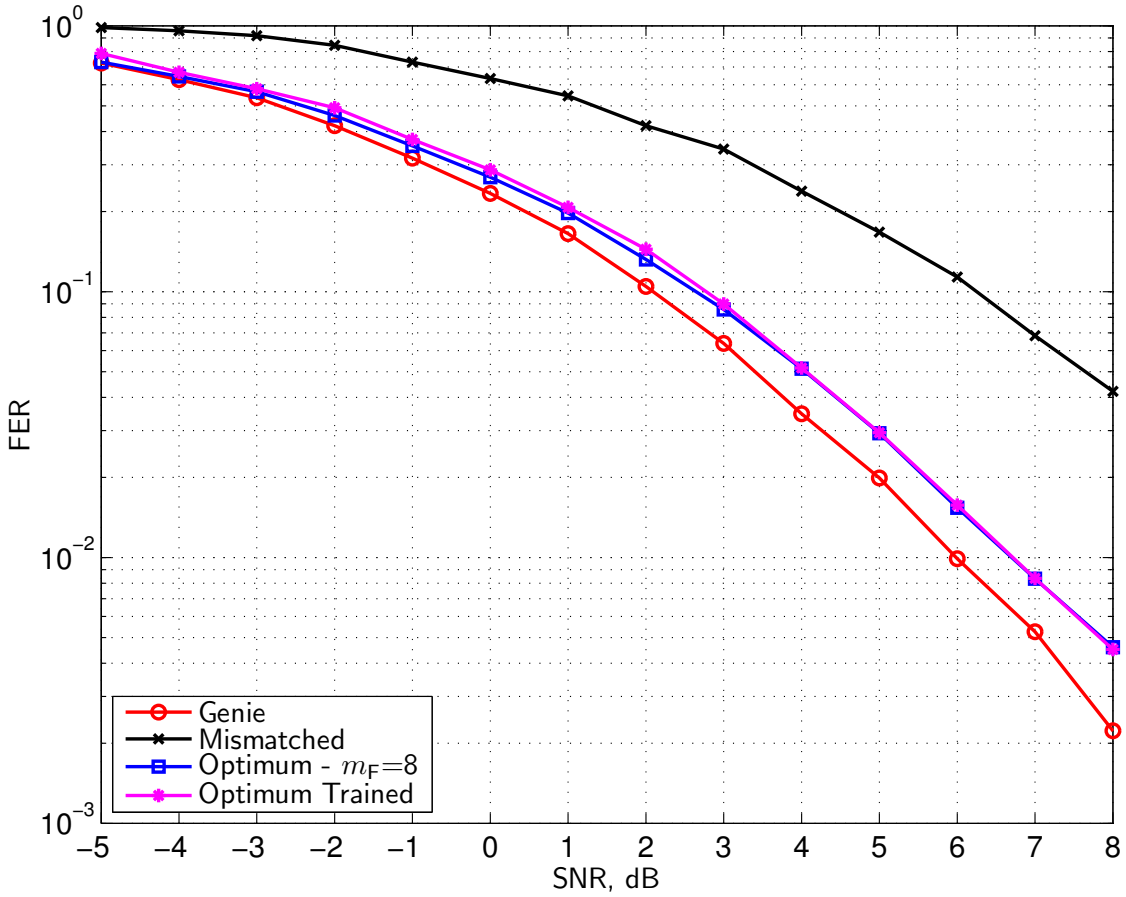


Figure 3.7. FER vs. SNR of the MIMO-OFDM system with $K = 0$ dB (Rice fading), $\rho_R = \rho_T = 0.7$. Genie, mismatched, optimum and optimum trained receivers.

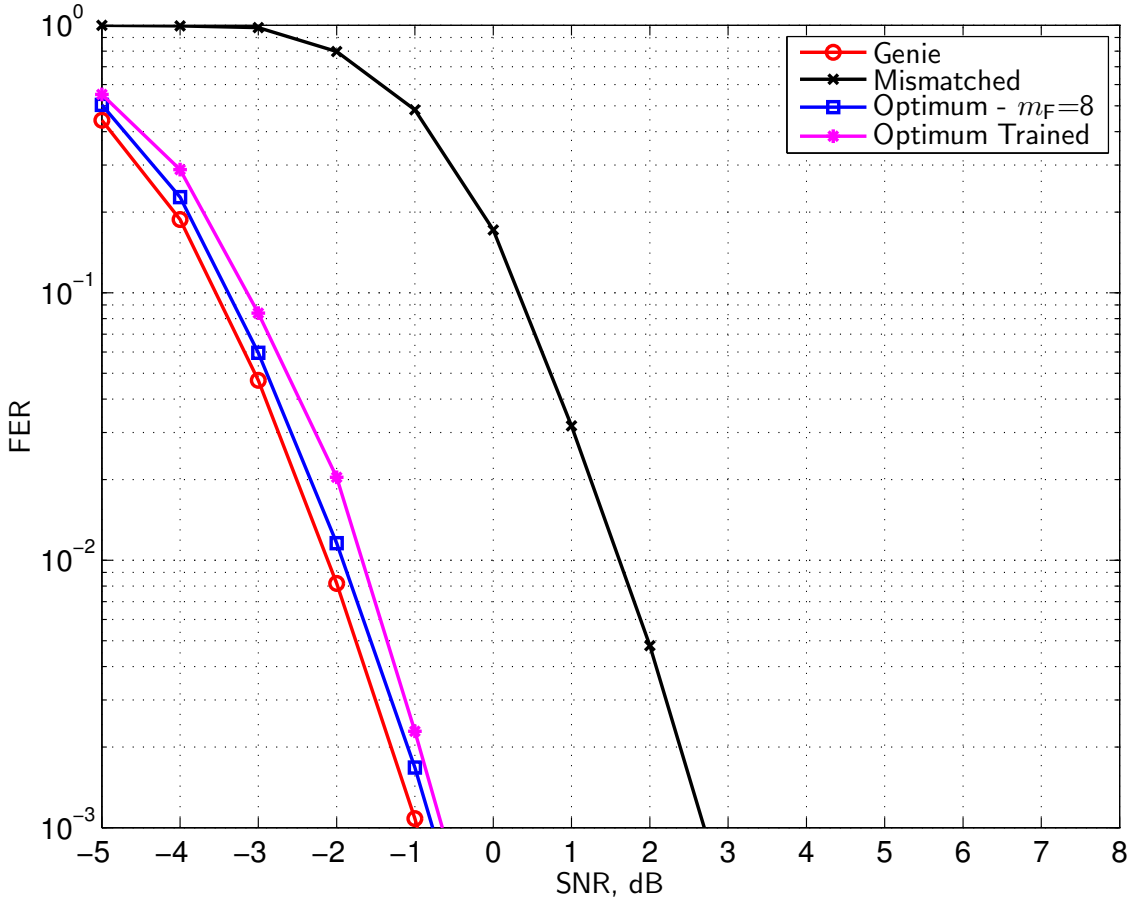


Figure 3.8. FER vs. SNR of the MIMO-OFDM system with $K = 10$ dB (Rice fading), $\rho_R = \rho_T = 0.0$. Genie, mismatched, optimum and optimum trained receivers.

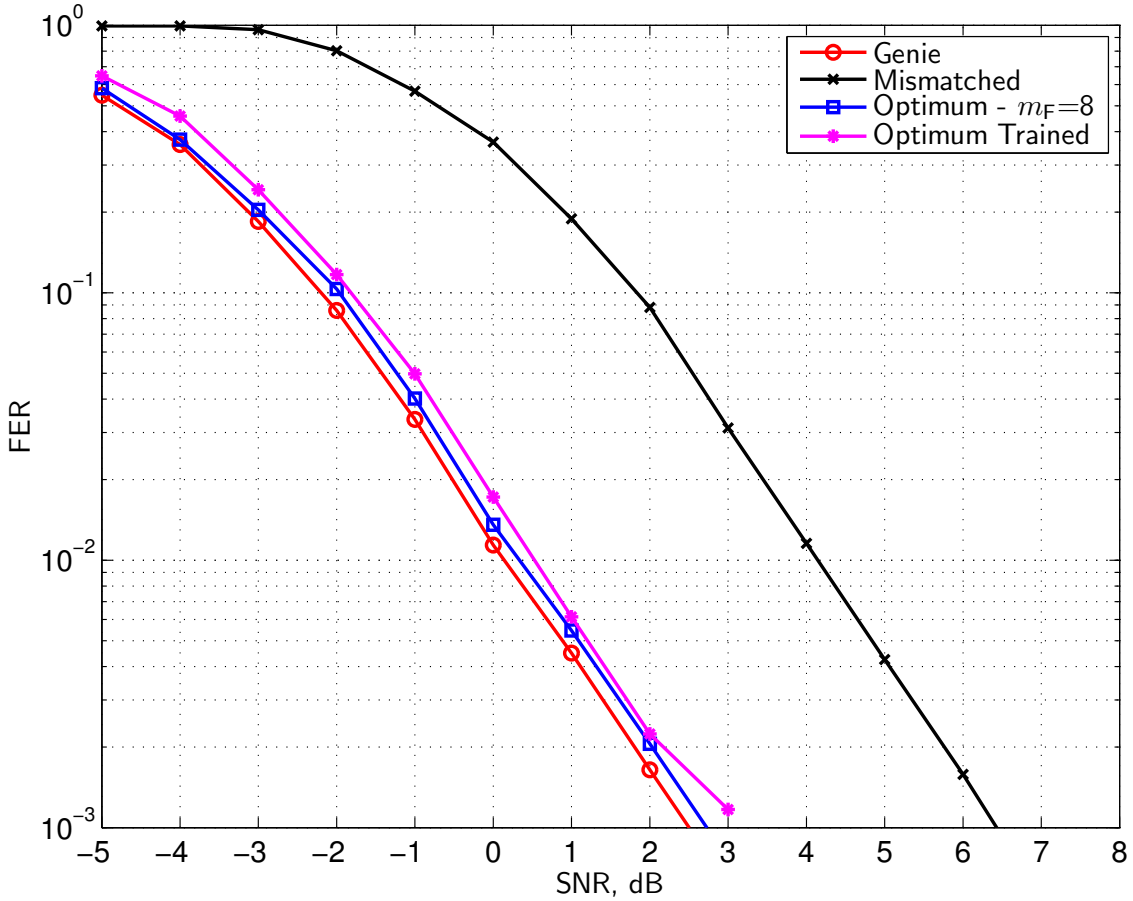


Figure 3.9. FER vs. SNR of the MIMO-OFDM system with $K = 10$ dB (Rice fading), $\rho_R = \rho_T = 0.7$. Genie, mismatched, optimum and optimum trained receivers.

3.3.4 Trained receiver: effect of imperfect knowledge of channel parameters

The numerical results reported in Figs. 3.6 to 3.9 refer to the three main receivers considered in this work (mismatched, optimum, and genie) as well as an additional fourth receiver referred to as optimum trained receiver. The optimum trained receiver implements the parameter estimation algorithm developed in Section 3.2.7 (with a Training mode phase of 10 frames) to derive the channel parameters.

It is worth to underline the fact that information symbols are also transmitted during the training phase so that its occurrence does not reduce the system throughput as one might think at first sight because the parameter estimation algorithm runs concurrently with the decoding algorithm, which is switched to the mismatched receiver in the training phase and to the optimum receiver in the operating phase.

As the curves show, the performance loss of the optimum trained receiver with respect to the optimum receiver is always limited to a few tenths of one dB, so that it can be safely stated that, at least in the cases considered, the need for parameter estimation does not affect significantly the achievable error performance of the optimum receiver.

Chapter 4

Conclusions

In this thesis, two receiver architectures for coded MIMO systems based on pilot symbol-aided channel estimation and the effects of non ideal parameter estimation were studied.

The receiver structures are based on two different architectures:

1. a mismatched receiver, which detects by ML or MMSE estimation the MIMO channel matrix and decodes the received signal as if this estimate were perfect;

2. an optimum receiver, which skips the channel matrix detection phase and decodes the received signal by jointly processing the pilot and data samples.

It was shown that it is possible to obtain in a closed form an optimum decision metric, for both the considered channel models:

1. a narrowband channel model with separate spatial correlation at the transmitter and at the receiver;

2. a MIMO–OFDM channel model with separate spatial correlation at the transmitter and at the receiver and in the frequency domain among subcarriers.

Summarizing, the main results regarding the narrowband optimum receiver are listed as follows.

- After developing the detection algorithms (and in particular the branch metric increments for the application to trellis decoding), we obtained simulation results under several fading conditions.

- Numerical results relevant to the case of Rayleigh fading and joint transmit/receive correlation show that the optimum receiver gains at least 1 dB over the mismatched receiver with a throughput reduction (due to pilot symbol insertion) of approximately 3% for the former ($P = 4$) and 11% for the

latter ($P = 16$). Fixing the throughput reduction to 3% in both cases, the gain reaches approximately equal to 3 dB.

- On the other hand, it must be recognized that the higher complexity of the optimum receiver may be an issue in high-speed or tight-budget implementations but the continuous development of faster and inexpensive signal processing devices will make this issue negligible in a convenient time frame.

- Moreover, it was showed that the performance degradation due to non ideal parameter estimation (based on a simple suboptimum algorithm) is almost unnoticeable and the optimum receiver performance is achieved after a very short training period.

- For this reason the optimum receiver, based on pilot-aided channel estimation, was tested over *measured* MIMO channels in urban and rich scattering environments.

- Numerical results show that the proposed optimum receiver loses approximately 0.4 dB when compared to the genie receiver (based on perfect CSI) while the standard *mismatched* receiver may lose over 3 dB (at maximum throughput).

- Finally, a blind receiver scheme for coded MIMO systems was proposed and compared in terms of error performance to that of the optimum pilot-aided receiver.

- It was showed that, when the LOS channel component is sufficiently strong, the proposed blind receiver outperforms the pilot-aided receiver in all cases. On the contrary, as the Rice factor gets smaller, the blind receiver outperforms the pilot-aided receiver only when the E_b/N_0 ratio exceeds a specific threshold.It is worth noting that the blind receiver achieves the maximum throughput while the pilot-aided receiver throughput is bounded below the maximum.

- Moreover, the blind receiver complexity is smaller as it does not need to distinguish between pilot and data sections, which represents an advantage for limited-complexity devices such as those used for wireless sensor networks.

To sum up the results obtained for the optimum MIMO–OFDM receiver, the following considerations can be stated.

- The receiver architecture is modeled around the emerging IEEE 802.11n standard encompassing MIMO transmission. The receiver resorts to pilot-aided channel estimation in a different way than other receivers as it performs jointly

channel estimation and data detection. Though based on perfect channel distribution information, the receiver has been shown to be only slightly affected by imperfect parameter estimation accomplished by a very simple algorithm running in parallel with the main decoding algorithm.

- The novel optimum receiver has been studied in detailed for the MIMO-OFDM system as far as concerns its metric in the presence of spatially correlated Rician fading.
- Complexity issues have been addressed in two separate directions:
  - by implementing the metric computation iteratively;
  - by implementing a spectral approximation technique which allows a dramatic reduction of the decoding overhead.
- Error performance results have shown that the optimum receiver closes the gap to the genie receiver within 0.5–0.7 dB while the mismatched receiver loses more than 4 dB in the cases considered (Rayleigh and Rician fading, uncorrelated or strongly correlated).
- Finally, parameter estimation, required to the optimum receiver, has been shown not to affect significantly the achievable error performance and to account for an approximate loss of a few tenths of one dB.

Appendix A

Mathematical details for Optimum Narrowband Receiver

A.1 Proof of Lemma (2.1.1)

Proof. The covariance can be evaluated as follows:

$$
\begin{aligned}
&\operatorname{cov}\left((\mathbf{H})_{ij},(\mathbf{H})_{i'j'}\right)\\
=&\mathbb{E}\left[(\mathbf{H})_{ij}(\mathbf{H})^*_{i'j'}\right]-\mathbb{E}\left[(\mathbf{H})_{ij}\right]\mathbb{E}\left[(\mathbf{H})_{i'j'}\right]^*\\
=&\mathbb{E}\left[\sqrt{\frac{1}{K+1}}\left(\mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}\right)_{ij}\sqrt{\frac{1}{K+1}}\left(\mathbf{R}^{1/2}\mathbf{W}\mathbf{T}^{1/2}\right)_{i'j'}\right]\\
=&\frac{1}{K+1}\mathbb{E}\Bigg[\sum_{a=1}^{n_{\mathsf{R}}}\sum_{b=1}^{n_{\mathsf{T}}}\sum_{a'=1}^{n_{\mathsf{R}}}\sum_{b'=1}^{n_{\mathsf{T}}}\left(\mathbf{R}^{1/2}\right)_{ia}\left(\mathbf{W}\right)_{ab}\left(\mathbf{T}^{1/2}\right)_{bj}\\
&\cdot\left(\mathbf{R}^{1/2}\right)^*_{i'a'}\left(\mathbf{W}\right)^*_{a'b'}\left(\mathbf{T}^{1/2}\right)^*_{b'j'}\Bigg]\\
=&\frac{1}{K+1}\sum_{a=1}^{n_{\mathsf{R}}}\sum_{b=1}^{n_{\mathsf{T}}}\left(\mathbf{R}^{1/2}\right)_{ia}\left(\mathbf{T}^{1/2}\right)_{bj}\left(\mathbf{R}^{1/2}\right)^*_{i'a}\left(\mathbf{T}^{1/2}\right)^*_{bj'}\\
=&\frac{1}{K+1}\sum_{a=1}^{n_{\mathsf{R}}}\left(\mathbf{R}^{1/2}\right)_{ia}\left(\mathbf{R}^{1/2}\right)_{ai'}\sum_{b=1}^{n_{\mathsf{T}}}\left(\mathbf{T}^{1/2}\right)^*_{jb}\left(\mathbf{T}^{1/2}\right)^*_{bj'}\\
=&\frac{1}{K+1}(\mathbf{R})_{ii'}(\mathbf{T})^*_{jj'}
\end{aligned}
$$

□

A.2 Optimum metric derivation

The following result was used in the derivation of the optimum receiver metric.

Theorem A.2.1. *Let* $\mathbf{H}$ *be an* $m \times n$ *matrix of independent circularly-symmetric zero-mean complex Gaussian random variables with unit variance. Let* $\mathbf{A}$ *and* $\mathbf{R}$ *be Hermitian matrices of dimensions* $n \times n$ *and* $m \times m$*, respectively, and* $\mathbf{B}$ *be any* $m \times n$ *matrix. Then, the following identity holds:*

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}}\left[\mathrm{etr}\left(-(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}+\mathbf{H}\mathbf{B}^{\mathsf{H}}+\mathbf{B}\mathbf{H}^{\mathsf{H}})\right)\right]\\&=\det(\mathbf{I}_{mn}+\mathbf{C})^{-1}\exp\left(\mathrm{vec}\left\{\mathbf{B}\right\}^{\mathsf{H}}(\mathbf{I}_{mn}+\mathbf{C})^{-1}\mathrm{vec}\left\{\mathbf{B}\right\}\right)\end{aligned} \tag{A.1}$$

where $\mathbf{C} \triangleq \mathbf{A}^{\mathsf{T}} \otimes \mathbf{R}$.

Proof. The expectation can be written as follows:

$$\begin{aligned}&\mathbb{E}_{\mathbf{H}}\left[\mathrm{etr}\left(-(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}+\mathbf{H}\mathbf{B}^{\mathsf{H}}+\mathbf{B}\mathbf{H}^{\mathsf{H}})\right)\right]\\&=\int_{\mathbb{C}^{m\times n}}\pi^{-mn}\mathrm{etr}\left(-\mathbf{H}\mathbf{H}^{\mathsf{H}}\right)\mathrm{etr}\left(-(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}+\mathbf{H}\mathbf{B}^{\mathsf{H}}+\mathbf{B}\mathbf{H}^{\mathsf{H}})\right)\mathrm{d}\mathbf{H}\end{aligned} \tag{A.2}$$

Then, the term $\mathrm{Tr}\left(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}\right)$ is expanded as follows:

$$\mathrm{Tr}\left(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}\right)=\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{m}(\mathbf{H})^{*}_{\ell k}(\mathbf{A})_{jk}(\mathbf{R})_{\ell i}(\mathbf{H})_{ij}\ .$$

It can be noticed that this expression is equivalent to

$$\mathrm{Tr}\left(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}\right)=\mathbf{h}^{\mathsf{H}}\mathbf{C}\mathbf{h}$$

where

$$\mathbf{h} \triangleq \mathrm{vec}\left\{\mathbf{H}\right\}$$

and

$$\mathbf{C}=\mathbf{A}^{\mathsf{T}}\otimes\mathbf{R}\ .$$

$\mathbf{C}$ is Hermitian since

$$\mathbf{C}^{\mathsf{H}}=(\mathbf{A}^{\mathsf{T}})^{\mathsf{H}}\otimes\mathbf{R}^{\mathsf{H}}=\mathbf{A}^{\mathsf{T}}\otimes\mathbf{R}\ .$$

Therefore, from (A.2) teh following result can be obtained:

$$\begin{aligned}
&\mathbb{E}_{\mathbf{H}}\left[\operatorname{etr}\left(-(\mathbf{H}\mathbf{A}\mathbf{H}^{\mathsf{H}}\mathbf{R}+\mathbf{H}\mathbf{B}^{\mathsf{H}}+\mathbf{B}\mathbf{H}^{\mathsf{H}})\right)\right]\\
=&\pi^{-mn}\int_{\mathbb{C}^{mn}}\exp\left(-(\mathbf{h}^{\mathsf{H}}(\mathbf{I}_{mn}+\mathbf{C})\mathbf{h}+\mathbf{h}^{\mathsf{H}}\mathbf{b}+\mathbf{b}^{\mathsf{H}}\mathbf{h})\right)\mathrm{d}\mathbf{h}\\
=&\pi^{-mn}\exp\left(\mathbf{b}^{\mathsf{H}}(\mathbf{I}_{mn}+\mathbf{C})^{-1}\mathbf{b}\right)\\
&\int_{\mathbb{C}^{mn}}\exp\left(-\|(\mathbf{I}_{mn}+\mathbf{C})^{1/2}\mathbf{h}+(\mathbf{I}_{mn}+\mathbf{C})^{-1/2}\mathbf{b}\|^2\right)\mathrm{d}\mathbf{h}\\
=&\pi^{-mn}\det(\mathbf{I}_{mn}+\mathbf{C})^{-1}\exp\left(\mathbf{b}^{\mathsf{H}}(\mathbf{I}_{mn}+\mathbf{C})^{-1}\mathbf{b}\right)\\
&\int_{\mathbb{C}^{mn}}\exp\left(-\|\mathbf{u}\|^2\right)\mathrm{d}\mathbf{u}\\
=&\det(\mathbf{I}_{mn}+\mathbf{C})^{-1}\exp\left(\mathbf{b}^{\mathsf{H}}(\mathbf{I}_{mn}+\mathbf{C})^{-1}\mathbf{b}\right)
\end{aligned}$$

after applying the change of variables

$$\mathbf{h}=(\mathbf{I}_{mn}+\mathbf{C})^{-1/2}\mathbf{u}-(\mathbf{I}_{mn}+\mathbf{C})^{-1}\mathbf{b},$$

whose Jacobian is $\det(\mathbf{I}_{mn}+\mathbf{C})^{-1}$ □

A.3 Derivation of branch metric (2.16)

Writing explicitly the dependency of the metric (2.16) on the code word $\mathbf{X}$, the following expression can be obtained:

$$\begin{aligned}\mu_{\mathsf{OR}}(\mathbf{X}) =&\|\mathbf{Y}-\bar{\mathbf{H}}_0\mathbf{X}\|^2 - N_0 \ln\det(\mathbf{\Lambda}(\mathbf{X}))\\ &- N_0 \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}^{\mathsf{H}} \mathbf{\Lambda}(\mathbf{X}) \mathrm{vec}\left\{\mathbf{B}(\mathbf{X})\right\}\end{aligned}$$

where the matrices $\mathbf{A}(\mathbf{X})$ and $\mathbf{B}(\mathbf{X})$ are given in (2.12-2.13) and $\mathbf{\Lambda}(\mathbf{X}) \triangleq (\mathbf{I}_{n_{\mathsf{T}} n_{\mathsf{R}}} + \mathbf{A}(\mathbf{X})^{\mathsf{T}} \otimes \mathbf{R})^{-1}$.

Then, the metric increment defined as

$$\Delta\mu_{\mathsf{OR}}(\mathbf{X}^-,\mathbf{x}) \triangleq \mu_{\mathsf{OR}}(\mathbf{X}) - \mu_{\mathsf{OR}}(\mathbf{X}^-)$$

has to be calculated.

To this purpose define

$$\Delta\mathbf{B}(\mathbf{x}) \triangleq \mathbf{B}(\mathbf{X}) - \mathbf{B}(\mathbf{X}^-) = \mathbf{R}^{1/2} \frac{\mathbf{y}-\bar{\mathbf{H}}_0\mathbf{x}}{N_0\sqrt{K+1}} \mathbf{x}^{\mathsf{H}} \mathbf{T}^{1/2}$$

and

$$\begin{aligned}\Delta\mathbf{\Lambda}(\mathbf{X}^-,\mathbf{x}) &\triangleq \mathbf{\Lambda}(\mathbf{X}) - \mathbf{\Lambda}(\mathbf{X}^-)\\ &= \left[\mathbf{\Lambda}(\mathbf{X}^-)^{-1} + \left(\frac{\mathbf{T}^{1/2}\mathbf{x}\mathbf{x}^{\mathsf{H}}\mathbf{T}^{1/2}}{N_0(K+1)}\right)^{\mathsf{T}} \otimes \mathbf{R}\right]^{-1} - \mathbf{\Lambda}(\mathbf{X}^-)\end{aligned}$$

To further simplify the previous expression resort to the matrix inversion lemma [55]:

$$(\mathbf{A}+\mathbf{BC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}[\mathbf{I}+\mathbf{C}\mathbf{A}^{-1}\mathbf{B}]^{-1}\mathbf{C}\mathbf{A}^{-1}$$

(holding for invertible matrices $\mathbf{A}$ and $(\mathbf{A}+\mathbf{BC})$) and to the standard property of the Kronecker product [55]:

$$(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A}\otimes\mathbf{C})(\mathbf{B}\otimes\mathbf{D}).$$

Thus, defining

$$\mathbf{\Psi}(\mathbf{x}) \triangleq \frac{(\mathbf{T}^{1/2}\mathbf{x})^*}{\sqrt{N_0(K+1)}} \otimes \mathbf{R}^{1/2} \ ,$$

so that

$$\mathbf{\Psi}(\mathbf{x})\mathbf{\Psi}(\mathbf{x})^{\mathsf{H}} = \left(\frac{\mathbf{T}^{1/2}\mathbf{x}\mathbf{x}^{\mathsf{H}}\mathbf{T}^{1/2}}{N_0(K+1)}\right)^{\mathsf{T}} \otimes \mathbf{R} \ ,$$

the following result can be obtained:

$$\Delta\boldsymbol{\Lambda}(\mathbf{X}^{-},\mathbf{x}) = -\boldsymbol{\Lambda}(\mathbf{X}^{-})\boldsymbol{\Psi}(\mathbf{x})[\mathbf{I}_{n_{\mathsf{R}}} + \boldsymbol{\Psi}(\mathbf{x})^{\mathsf{H}}\boldsymbol{\Lambda}(\mathbf{X}^{-})\boldsymbol{\Psi}(\mathbf{x})]^{-1}\boldsymbol{\Psi}(\mathbf{x})^{\mathsf{H}}\boldsymbol{\Lambda}(\mathbf{X}^{-})$$

Finally, using some additional simplification from standard linear algebra [55], an iterative expansion of the term $\ln\det\{\boldsymbol{\Lambda}(\mathbf{X})\}$ can be obtained:

$$\begin{aligned}
&\ln\det\{\boldsymbol{\Lambda}(\mathbf{X})\} \\
&= \ln\det\left\{\boldsymbol{\Lambda}(\mathbf{X}^{-}) - \boldsymbol{\Lambda}(\mathbf{X}^{-})\boldsymbol{\Psi}(\mathbf{x})^{\mathsf{H}}[\mathbf{I}_{n_{\mathsf{R}}} + \boldsymbol{\Psi}(\mathbf{x})\boldsymbol{\Lambda}(\mathbf{X}^{-})\boldsymbol{\Psi}(\mathbf{x})^{\mathsf{H}}]^{-1}\boldsymbol{\Psi}(\mathbf{x})\boldsymbol{\Lambda}(\mathbf{X}^{-})\right\} \\
&= \ln\det\{\boldsymbol{\Lambda}(\mathbf{X}^{-})\} - \ln\det\{\mathbf{I}_{n_{\mathsf{R}}} + \boldsymbol{\Psi}(\mathbf{x})\boldsymbol{\Lambda}(\mathbf{X}^{-})\boldsymbol{\Psi}(\mathbf{x})^{\mathsf{H}}\}
\end{aligned}$$

Gathering all previous results, the differential optimum metric (2.16) can be obtained.

Appendix B

Mathematical details for Optimum MIMO-OFDM receiver

B.1 Proof of Lemma (3.1.1)

Proof. The covariance can be evaluated as follows:

$$
\begin{aligned}
&\operatorname{cov}\left((\mathbf{H}_k)_{ij},(\mathbf{H}_{k'})_{i'j'}\right)\\
=&\mathbb{E}\left[(\mathbf{H}_k)_{ij}(\mathbf{H}_{k'})^*_{i'j'}\right]-\mathbb{E}\left[(\mathbf{H}_k)_{ij}\right]\mathbb{E}\left[(\mathbf{H}_{k'})_{i'j'}\right]^*\\
=&\mathbb{E}\left[\left(\left((\mathbf{F}^{1/2})_k\otimes\mathbf{R}^{1/2}\right)\mathbf{W}\mathbf{T}^{1/2}\right)_{ij}\left(\left((\mathbf{F}^{1/2})_{k'}\otimes\mathbf{R}^{1/2}\right)\mathbf{W}\mathbf{T}^{1/2}\right)_{i'j'}\right]\\
=&\mathbb{E}\Bigg[\sum_{a=1}^{n_{\mathsf{R}}n_{\mathsf{F}}}\sum_{b=1}^{n_{\mathsf{T}}}\sum_{a'=1}^{n_{\mathsf{R}}n_{\mathsf{F}}}\sum_{b'=1}^{n_{\mathsf{T}}}\left((\mathbf{F}^{1/2})_k\otimes\mathbf{R}^{1/2}\right)_{ia}\left(\mathbf{W}\right)_{ab}\left(\mathbf{T}^{1/2}\right)_{bj}\\
&\cdot\left((\mathbf{F}^{1/2})_{k'}\otimes\mathbf{R}^{1/2}\right)^*_{i'a'}\left(\mathbf{W}\right)^*_{a'b'}\left(\mathbf{T}^{1/2}\right)^*_{b'j'}\Bigg]\\
=&\sum_{a=1}^{n_{\mathsf{R}}n_{\mathsf{F}}}\sum_{b=1}^{n_{\mathsf{T}}}\left((\mathbf{F}^{1/2})_k\otimes\mathbf{R}^{1/2}\right)_{ia}\left(\mathbf{T}^{1/2}\right)_{bj}\left((\mathbf{F}^{1/2})_{k'}\otimes\mathbf{R}^{1/2}\right)^*_{i'a}\left(\mathbf{T}^{1/2}\right)^*_{bj'}\\
=&\sum_{a=1}^{n_{\mathsf{R}}n_{\mathsf{F}}}\left((\mathbf{F}^{1/2})_k\otimes\mathbf{R}^{1/2}\right)_{ia}\left((\mathbf{F}^{1/2})^{\mathsf{H}}_{k'}\otimes\mathbf{R}^{1/2}\right)_{ai'}\sum_{b=1}^{n_{\mathsf{T}}}\left(\mathbf{T}^{1/2}\right)^*_{jb}\left(\mathbf{T}^{1/2}\right)^*_{bj'}\\
=&\left(\left((\mathbf{F}^{1/2})_k\otimes\mathbf{R}^{1/2}\right)\left((\mathbf{F}^{1/2})^{\mathsf{H}}_{k'}\otimes\mathbf{R}^{1/2}\right)\right)_{ii'}\left(\mathbf{T}\right)^*_{jj'}\\
=&\left((\mathbf{F}^{1/2})_k(\mathbf{F}^{1/2})^{\mathsf{H}}_{k'}\otimes\mathbf{R}\right)_{ii'}\left(\mathbf{T}\right)^*_{jj'}\\
=&(\mathbf{R})_{ii'}(\mathbf{T})^*_{jj'}(\mathbf{F})_{kk'}
\end{aligned}
$$

since $(\mathbf{F}^{1/2})_k(\mathbf{F}^{1/2})^{\mathsf{H}}_{k'}=(\mathbf{F})_{kk'}$. □

B.2 Proof of Proposition 3.2.1

Proof. To simplifiy notation define the matrices

$$\widetilde{\mathbf{R}}_k \triangleq (\mathbf{F}^{1/2})_k \otimes \mathbf{R}^{1/2}$$

and write the probability density functions as:

$$\begin{aligned} & p\left(\mathbf{Y}|\mathbf{X},\mathbf{H}\right) p\left(\mathbf{Y}_{\mathsf{P}}|\mathbf{X}_{\mathsf{P}},\mathbf{H}\right) \\ =&(\pi N_0)^{-(N+P)n_{\mathsf{R}}n_{\mathsf{F}}} \exp\left(-\frac{\|\mathbf{Y}_{\mathsf{P}}-\bar{\mathbf{H}}\mathbf{X}_{\mathsf{P}}\|^2+\|\mathbf{Y}-\bar{\mathbf{H}}\mathbf{X}\|^2}{N_0}\right) \\ &\cdot \operatorname{etr}\left(\sum_{k=1}^{n_F}(-\mathbf{W}\mathbf{A}_k\mathbf{W}^{\mathsf{H}}\widetilde{\mathbf{R}}_k^{\mathsf{H}}\widetilde{\mathbf{R}}_k+\widetilde{\mathbf{R}}_k^{\mathsf{H}}\mathbf{B}_k\mathbf{W}^{\mathsf{H}}+\mathbf{W}\mathbf{B}_k^{\mathsf{H}}\widetilde{\mathbf{R}}_k)\right). \end{aligned}$$

Applying Identity B.2.1 gives

$$\begin{aligned} & \mathbb{E}\left[p\left(\mathbf{Y}|\mathbf{X},\mathbf{H}\right) p\left(\mathbf{Y}_{\mathsf{P}}|\mathbf{X}_{\mathsf{P}},\mathbf{H}\right)\right] \\ =&(\pi N_0)^{-(N+P)n_{\mathsf{R}}n_{\mathsf{F}}} \exp\left(-\frac{\|\mathbf{Y}_{\mathsf{P}}-\bar{\mathbf{H}}\mathbf{X}_{\mathsf{P}}\|^2+\|\mathbf{Y}-\bar{\mathbf{H}}\mathbf{X}\|^2}{N_0}\right) \\ &\cdot \det\left(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\sum_{k=1}^{n_F}(\mathbf{A}_k^{\mathsf{T}}\otimes(\widetilde{\mathbf{R}}_k^{\mathsf{H}}\widetilde{\mathbf{R}}_k)\right)^{-1} \\ &\cdot \exp\left(\operatorname{vec}\left\{\sum_{k=1}^{n_F}\widetilde{\mathbf{R}}_k^{\mathsf{H}}\mathbf{B}_k\right\}^{\mathsf{H}}(\mathbf{I}_{n_{\mathsf{T}}n_{\mathsf{R}}n_{\mathsf{F}}}+\sum_{k=1}^{n_F}\mathbf{A}_k^{\mathsf{T}}\otimes(\widetilde{\mathbf{R}}_k^{\mathsf{H}}\widetilde{\mathbf{R}}_k))^{-1}\operatorname{vec}\left\{\sum_{k=1}^{n_F}\widetilde{\mathbf{R}}_k^{\mathsf{H}}\mathbf{B}_k\right\}\right) \end{aligned}$$

Inserting for $\widetilde{\mathbf{R}}_k$ gives the desired result. □

Identity B.2.1. *Let* $\mathbf{W}$ *be a* $m \times n$ *matrix of independent circularly-symmetric zero-mean complex Gaussian random variables with unit variance. Let* $\mathbf{A}_k$ *and* $\mathbf{R}_k$ *be complex positive semidefinite matrices of dimension* $n\times n$ *and* $m\times m$*, respectively and* $\mathbf{B}_k$ *be complex* $m \times n$ *matrices, where* $k \in \{1,\ldots,n_F\}$*. Then, the following identity holds:*

$$\begin{aligned} & \mathbb{E}_{\mathbf{W}}\left[\operatorname{etr}\left(-\sum_{k=1}^{n_F}(\mathbf{W}\mathbf{A}_k\mathbf{W}^{\mathsf{H}}\mathbf{R}_k+\mathbf{W}\mathbf{B}_k^{\mathsf{H}}+\mathbf{B}_k\mathbf{W}^{\mathsf{H}})\right)\right] \\ =&\det\left(\mathbf{I}_{mn}+\sum_{k=1}^{n_F}\mathbf{A}_k^{\mathsf{T}}\otimes\mathbf{R}_k\right)^{-1} \\ &\cdot\exp\left(\operatorname{vec}\left\{\sum_{k=1}^{n_F}\mathbf{B}_k\right\}^{\mathsf{H}}\left(\mathbf{I}_{mn}+\sum_{k=1}^{n_F}\mathbf{A}_k^{\mathsf{T}}\otimes\mathbf{R}_k\right)^{-1}\operatorname{vec}\left\{\sum_{k=1}^{n_F}\mathbf{B}_k\right\}\right). \end{aligned}$$

Proof. This follows from

$$\mathrm{Tr}\left(\mathbf{W}\mathbf{A}_k\mathbf{W}^{\mathsf{H}}\mathbf{R}_k\right) = \mathrm{vec}\left\{\mathbf{W}\right\}^{\mathsf{H}}\mathbf{A}_k^{\mathsf{T}} \otimes \mathbf{R}_k \mathrm{vec}\left\{\mathbf{W}\right\}$$

and [51, Theorem 1]. □

B.3 Proof of Proposition 3.2.2

First, proof the following technical lemma:

Lemma B.3.1. Define the matrices

$$\mathbf{C}_m \triangleq \mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \sum_{k=1}^{m} \mathbf{U}_k^\mathsf{H} \mathbf{U}_k$$

for $m = 1, \ldots, n_\mathsf{F}$.

Then

$$\mathbf{C}_m^{-1} = \mathbf{\Lambda}_m.$$

Proof. The proof is based on induction. From the matrix inversion Lemma [55, p.18] it is plain to see that

$$\mathbf{\Lambda}_k = \mathbf{\Lambda}_{k-1} + \Delta\mathbf{\Lambda} = (\mathbf{\Lambda}_{k-1}^{-1} + \mathbf{U}_k^\mathsf{H} \mathbf{U}_k)^{-1} \ .$$

Therefore,

$$\mathbf{\Lambda}_1 = (\mathbf{\Lambda}_0^{-1} + \mathbf{U}_1^\mathsf{H} \mathbf{U}_1)^{-1} = \mathbf{C}_1^{-1} \ .$$

For $m > 1$:

$$\begin{aligned}
\mathbf{C}_m^{-1} =& \mathbf{C}_{m-1}^{-1} (\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{U}_m^\mathsf{H} \mathbf{U}_m \mathbf{C}_{m-1}^{-1})^{-1} \\
=& \mathbf{\Lambda}_{m-1} (\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}} + \mathbf{U}_m^\mathsf{H} \mathbf{U}_m \mathbf{\Lambda}_{m-1})^{-1} \\
=& (\mathbf{\Lambda}_{m-1}^{-1} + \mathbf{U}_m^\mathsf{H} \mathbf{U}_m)^{-1} \\
=& \mathbf{\Lambda}_m.
\end{aligned}$$

□

Now return to the proof of Proposition 3.2.2 which is also based on induction:

Proof. Let

$$\mu_{\mathsf{OR}}^{(m)}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_\mathsf{P},\mathbf{X}_\mathsf{P}) \triangleq \sum_{k=1}^{m} \|\mathbf{Y}_k - \bar{\mathbf{H}}_k \mathbf{X}_k\|^2 + N_0 \ln\det(\mathbf{C}_m) - N_0 \mathbf{b}_m^\mathsf{H} \mathbf{C}_m^{-1} \mathbf{b}_m$$

for $m \leq n_\mathsf{F}$. Obviously $\mu_{\mathsf{OR}} = \mu_{\mathsf{OR}}^{(n_\mathsf{F})}$.

1. For $m = 1$:

$$\begin{aligned}
&\mu_{\mathsf{OR}}^{(1)}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_\mathsf{P},\mathbf{X}_\mathsf{P}) \\
=& \|\mathbf{Y}_1 - \bar{\mathbf{H}}_1 \mathbf{X}_1\|^2 + N_0 \ln\det(\mathbf{C}_1) - N_0 \mathbf{b}_1^\mathsf{H} \mathbf{C}_1^{-1} \mathbf{b}_1 \\
=& \|\mathbf{Y}_1 - \bar{\mathbf{H}}_1 \mathbf{X}_1\|^2 + N_0 \ln\det(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \mathbf{U}_1 \mathbf{\Lambda}_0 \mathbf{U}_1^\mathsf{H}) - N_0 \mathbf{b}_1^\mathsf{H} \mathbf{\Lambda}_1 \mathbf{b}_1 + N_0 \mathbf{b}_0^\mathsf{H} \mathbf{\Lambda}_0 \mathbf{b}_0
\end{aligned}$$

where the following properties were used

(a) $\mathbf{b}_0 = \mathbf{0}$,

(b) $\mathbf{\Lambda}_0 = \mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}}$,

(c) $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$ [55, p.53] and

(d) Lemma B.3.1.

2. Similarly, for $m > 1$:

$$
\begin{aligned}
&\mu_{\mathsf{OR}}^{(m)}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_\mathsf{P},\mathbf{X}_\mathsf{P}) \\
=&\sum_{k=1}^{m} \|\mathbf{Y}_k - \bar{\mathbf{H}}_k \mathbf{X}_k\|^2 + N_0 \ln\det(\mathbf{C}_{m-1} + \mathbf{U}_m^\mathsf{H}\mathbf{U}_m) - N_0 \mathbf{b}_m^\mathsf{H} \mathbf{C}_m^{-1} \mathbf{b}_m \\
=&\mu_{\mathsf{OR}}^{(m-1)}(\mathbf{X}|\mathbf{Y},\mathbf{Y}_\mathsf{P},\mathbf{X}_\mathsf{P}) + \|\mathbf{Y}_m - \bar{\mathbf{H}}_m \mathbf{X}_m\|^2 + N_0 \ln\det(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \mathbf{U}_m \mathbf{C}_{m-1}^{-1} \mathbf{U}_m^\mathsf{H}) \\
&- N_0 \mathbf{b}_m^\mathsf{H} \mathbf{C}_m^{-1} \mathbf{b}_m + N_0 \mathbf{b}_{m-1}^\mathsf{H} \mathbf{C}_{m-1}^{-1} \mathbf{b}_{m-1} \\
=&\sum_{k=1}^{m-1} \Delta\mu_k(\mathbf{X}_k,\mathbf{Y}_k,\mathbf{X}_{\mathsf{P},k},\mathbf{Y}_{\mathsf{P},k}) + \|\mathbf{Y}_m - \bar{\mathbf{H}}_m \mathbf{X}_m\|^2 + N_0 \ln\det(\mathbf{I}_{n_\mathsf{T} n_\mathsf{R}} + \mathbf{U}_m \mathbf{C}_{m-1}^{-1} \mathbf{U}_m^\mathsf{H}) \\
&- N_0 \mathbf{b}_m^\mathsf{H} \mathbf{C}_m^{-1} \mathbf{b}_m + N_0 \mathbf{b}_{m-1}^\mathsf{H} \mathbf{C}_{m-1}^{-1} \mathbf{b}_{m-1} \\
=&\sum_{k=1}^{m} \Delta\mu_k(\mathbf{X}_k,\mathbf{Y}_k,\mathbf{X}_{\mathsf{P},k},\mathbf{Y}_{\mathsf{P},k}) \ ,
\end{aligned}
$$

where in the last step Lemma B.3.1 was used.

□

B.4 Proof of Proposition 3.2.3

Proof. To simplify the notation, $\mathbf{I}$ is a shortcut for $\mathbf{I}_{n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}}$. It is plain to see that the metric expression (3.19) is invariant under the transformation $\mathbf{F}^{1/2} \rightarrow \mathbf{F}^{1/2}\mathbf{V} = \mathbf{V}\mathbf{D}$ which implies that $\mu_{\mathsf{OR}} = \|\mathbf{Y} - \bar{\mathbf{H}}\mathbf{X}\|^2 + \mu^{(1)}_{\mathsf{OR}} - \mu^{(2)}_{\mathsf{OR}}$.

Noting that

$$\mathbf{g}_k^{m_\mathsf{F}\mathsf{H}}\mathbf{g}_l^{m_\mathsf{F}} = (\mathbf{V}\mathbf{D}^2_{m_\mathsf{F}}\mathbf{V}^\mathsf{H})_{kl} \quad \overline{\mathbf{g}}_k^{m_\mathsf{F}\mathsf{H}}\overline{\mathbf{g}}_l^{m_\mathsf{F}} = (\mathbf{V}\overline{\mathbf{D}}^2_{m_\mathsf{F}}\mathbf{V}^\mathsf{H})_{kl} \quad \mathbf{g}_k^{m_\mathsf{F}\mathsf{H}}\overline{\mathbf{g}}_l^{m_\mathsf{F}} = \overline{\mathbf{g}}_k^{m_\mathsf{F}\mathsf{H}}\mathbf{g}_l^{m_\mathsf{F}} = 0 \tag{B.1}$$

it can be obtained:

$$\begin{aligned}
\|\overline{\mathbf{b}}_{m_\mathsf{F}}\|^2 =& \sum_{k,l=1}^{n_\mathsf{F}} \overline{\mathbf{g}}_k^{m_\mathsf{F}\mathsf{H}}\overline{\mathbf{g}}_l^{m_\mathsf{F}} \mathrm{Tr}\left(\mathbf{B}_k^\mathsf{H}\mathbf{R}\mathbf{B}_l\right) \leq \sqrt{\sum_{k,l=1}^{n_\mathsf{F}} |\overline{\mathbf{g}}_k^{m_\mathsf{F}\mathsf{H}}\overline{\mathbf{g}}_l^{m_\mathsf{F}}|^2} \sqrt{\sum_{k,l=1}^{n_\mathsf{F}} |\mathrm{Tr}\left(\mathbf{B}_k^\mathsf{H}\mathbf{R}\mathbf{B}_l\right)|^2} \\
\leq& \sqrt{n_\mathsf{F}} f_{m_\mathsf{F}+1} \sqrt{\sum_{k,l=1}^{n_\mathsf{F}} |\mathrm{Tr}\left(\mathbf{B}_k^\mathsf{H}\mathbf{R}\mathbf{B}_l\right)|^2}. \\
|||\overline{\mathbf{C}}_{m_\mathsf{F}}||| \leq& |||\mathbf{R}||| \cdot \sum_{k=1}^{n_\mathsf{F}} |||\mathbf{A}_k|||(\|\overline{\mathbf{g}}_k^{m_\mathsf{F}}\overline{\mathbf{g}}_k^{m_\mathsf{F}\mathsf{H}}\| + 2\|\overline{\mathbf{g}}_k^{m_\mathsf{F}}\mathbf{g}_k^{m_\mathsf{F}\mathsf{H}}\|) \\
\leq& \left(f_{m_\mathsf{F}+1} + 2\sqrt{|||\mathbf{F}||| \cdot f_{m_\mathsf{F}+1}}\right) \cdot |||\mathbf{R}||| \cdot \sum_{k=1}^{n_\mathsf{F}} |||\mathbf{A}_k|||
\end{aligned} \tag{B.2}$$

which gives equations (3.23).

Using the fact that

1. $\mathbf{b}_{m_\mathsf{F}}^\mathsf{H}\overline{\mathbf{b}}_{m_\mathsf{F}} = 0$ and $\mathbf{C}_{m_\mathsf{F}}\overline{\mathbf{b}}_{m_\mathsf{F}} = \mathbf{0}$,

2. $(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I} - (\mathbf{I}+\mathbf{A})^{-1}\mathbf{A}$ for any positive semidefinite $n_\mathsf{T} n_\mathsf{R} n_\mathsf{F} \times n_\mathsf{T} n_\mathsf{R} n_\mathsf{F}$ matrix $\mathbf{A}$,

3. Cauchy-Schwarz inequality [55, p.15],

4. sub-multiplicativity of $|||.|||$ and

5. $|||(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_\mathsf{F}}(\mathbf{I}+\mathbf{C})^{-1/2}||| \leq |||\overline{\mathbf{C}}_{m_\mathsf{F}}||| \cdot |||(\mathbf{I}+\mathbf{C})^{-1}||| < 1$

it can be obtained:

$$
\begin{aligned}
&|\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1}\mathbf{b}-\mathbf{b}_{m_{\mathsf{F}}}^{\mathsf{H}}(\mathbf{I}+\mathbf{C}_{m_{\mathsf{F}}})^{-1}\mathbf{b}_{m_{\mathsf{F}}}|\\
=&|\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1}\mathbf{b}-\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C}_{m_{\mathsf{F}}})^{-1}\mathbf{b}+\overline{\mathbf{b}}_{m_{\mathsf{F}}}{}^{\mathsf{H}}\overline{\mathbf{b}}_{m_{\mathsf{F}}}|\\
\leq&|\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1}\mathbf{b}-\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C}_{m_{\mathsf{F}}})^{-1}\mathbf{b}|+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
=&|\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1}\mathbf{b}-\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1/2}\sum_{k=0}^{\infty}\left((\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}\right)^k(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}|\\
&+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
=&|\mathbf{b}^{\mathsf{H}}(\mathbf{I}+\mathbf{C})^{-1/2}\sum_{k=1}^{\infty}\left((\mathbf{I}+\mathbf{C})^{-1/2})\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}\right)^k(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}|+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
\leq&\|(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}\|^2\cdot|||\sum_{k=1}^{\infty}\left((\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}\right)^k|||+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
=&\|(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}\|^2\cdot|||\sum_{k=1}^{\infty}\left((\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}\right)^k|||+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
\leq&\|(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}\|^2\cdot\sum_{k=1}^{\infty}\left(|||(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}|||\right)^k+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2\\
=&\|(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{b}\|^2\cdot\frac{|||(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}|||}{1-|||(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2}|||}+\|\overline{\mathbf{b}}_{m_{\mathsf{F}}}\|^2. \qquad \text{(B.3)}
\end{aligned}
$$

The approximation error for $\widetilde{\mu}_{\mathsf{OR}}^{(1)}$ is bounded by:

$$
\begin{aligned}
|\ln\det(\mathbf{I}+\mathbf{C}_{m_{\mathsf{F}}})-\ln\det(\mathbf{I}+\mathbf{C})| =&|\ln\det(\mathbf{I}-(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2})|\\
\leq&n_{\mathsf{F}}|\ln(1-|||(\mathbf{I}+\mathbf{C})^{-1/2}\overline{\mathbf{C}}_{m_{\mathsf{F}}}(\mathbf{I}+\mathbf{C})^{-1/2})|||)|. \qquad \text{(B.4)}
\end{aligned}
$$

Combining equation (B.4) and (B.3) and using the sub-multiplicativity of $|||.|||$ gives equations (3.22), which completes the proof. □

Bibliography

[1] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[2] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.

[3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, November - December 1999.

[4] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[5] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[6] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, July 1999.

[7] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "MIMO wireless channels: capacity and performance prediction," *Global Telecommunications Conference 2000*, vol. 2, pp. 1083–1088, November 27 - December 1 2000.

[8] ——, "Outdoor MIMO wireless channels: models and performance prediction," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 1926–1934, Dec 2002.

[9] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of blast in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, November 2000.

[10] D. Shiu, G. Foschini, M. Gans, J. Kahn, Q. Inc, and C. Santa Clara, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, March

2000.

[11] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, September 2001.

[12] C. Martin and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1350–1359, July 2004.

[13] S. Jayaweera and H. Poor, "MIMO capacity results for Rician fading channels," *IEEE Global Telecommunications Conference 2003*, vol. 4, pp. 1806–1810 vol.4, December 2003.

[14] L. Cottatellucci and M. Debbah, "On the capacity of MIMO Rice channels," *Proceedings of the 42nd Allerton Conference*, September 29 – October 1 2004.

[15] S. Venkatesan, S. Simon, and R. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," *Proc. IEEE Vehicular Technology Conference 2003*, vol. 3, pp. 1767–1771, October 2003.

[16] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[17] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, July 2004.

[18] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[19] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple-input-multiple-output measurements and modeling in Manhattan," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 321–331, April 2003.

[20] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, August 2002.

[21] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, "Deficiencies of 'Kronecker' MIMO radio channel model," *Electronics Letters*, vol. 39, no. 16, pp. 1209–1210, August 2003.

[22] B. Maharaj, L. Linde, and J. Wallace, "MIMO channel modelling: The Kronecker model and Maximum Entropy," *IEEE Wireless Communications and Networking Conference*, pp. 1909–1912, March 2007.

[23] R. Muller, "A random matrix model of communication via antenna arrays," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2495–2506, September 2002.

[24] A. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, October 2002.

[25] A. Tulino, A. Lozano, and S. Verdu, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 662–671, March 2006.

[26] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–100, January 2006.

[27] J. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, Nov 1991.

[28] J. Cavers and P. Ho, "Analysis of the error performance of trellis-coded modulations in Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 40, no. 1, pp. 74–83, January 1992.

[29] T. Marzetta, "BLAST training: Estimating channel characteristics for high capacity space-time wireless," *37th Annual Allerton Conference om Communication Control and Computing*, vol. 37, pp. 958–966, 1999.

[30] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, April 2003.

[31] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[32] L. Zheng and D. Tse, "Communication on the Grassmann manifold: a geometric approach tothe noncoherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, February 2002.

[33] M. McKay and I. Collings, "General Capacity Bounds for Spatially Correlated Rician MIMO Channels," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3121–3145, September 2005.

[34] X. Cui, Q. Zhang, and Z. Feng, "Generic procedure for tightly bounding the capacity of MIMO correlated Rician fading channels," *IEEE Transactions on Communications*, vol. 53, no. 5, pp. 890–898, May 2005.

[35] G. Taricco, "On the capacity of separately-correlated MIMO Rician fading channels," *Proc. IEEE Globecom 2006*, pp. 1–5, November 27 - December 1 2006.

[36] G. Taricco and E. Biglieri, "Space-time decoding with imperfect channel estimation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1874–1888, July 2005.

[37] V. Tarokh, A. Naguib, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of

channel estimation errors, mobility, and multiple paths," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 199–207, February 1999.

[38] ——, "Errata to 'Space-time codes for high data rate wireless communications: performance criteria in the presence of channel estimation errors, mobility, and multiple paths'," *IEEE Transactions on Communications*, vol. 51, no. 12, p. 2141, December 2003.

[39] A. Wiesel, Y. Eldar, and A. Beck, "Maximum Likelihood Estimation in Linear Models With a Gaussian Model Matrix," *IEEE Signal Processing Letters*, vol. 13, no. 5, p. 292, May 2006.

[40] K. Lee and J. Chun, "Symbol detection in V-BLAST architectures under channel estimation errors," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 593–597, Febraury 2007.

[41] H. Hofstetter, C. Mecklenbrauker, R. Muller, H. Anegg, H. Kunczier, E. Bonek, I. Viering, and A. Molisch, "Description of wireless MIMO measurements at 2 GHz in selected environments," *COST-273 5th Meeting, TD (02)*, vol. 135, September 18–19 2002.

[42] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communicationsin Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 543–564, March 2000.

[43] W.-K. Ma, B.-N. Vo, T. Davidson, and P.-C. Ching, "Blind ml detection of orthogonal space-time block codes: efficient high-performance implementations," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 738–751, February 2006.

[44] D. Guo and X. Wang, "Blind detection in MIMO systems via sequential Monte Carlo," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 464–473, April 2003.

[45] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 651–668, March 2002.

[46] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, October 1998.

[47] M. Brehler and M. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh-fading channels with applications to a unified analysis ofcoherent and noncoherent space-time receivers," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2383–2399, September 2001.

[48] J. Belfiore and A. Cipriano, "Space-time coding for non-coherent channels," *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, 2006.

[49] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, IEEE 802.11n Draft STANDARD for Information Technology, 2007.

[50] G. Taricco and G. Coluccia, "Mismatched and Optimum Receivers for the Correlated Rician Fading MIMO Channel," in *Proc. 44th Allerton Conference on Communication, Control, and Computing, September*, 2006.

[51] ——, "Optimum Receiver Design for Correlated Rician Fading MIMO Channels with Pilot-Aided Detection," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1311–1321, September 2007.

[52] G. Coluccia, G. Taricco, and C. Mecklenbräuker, "Performance of an optimum receiver scheme based on pilot-symbol channel estimation over a measured MIMO channel," *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications 2007*, pp. 1–5, June 2007.

[53] G. Coluccia and G. Taricco, "An Optimum Blind Receiver for Correlated Rician Fading MIMO Channels," *Communications Letters, IEEE*, vol. 11, no. 9, pp. 738–739, September 2007.

[54] G. Coluccia, E. Riegler, C. Mecklenbräuker, and G. Taricco, "Optimum MIMO-OFDM Detection with Pilot-Aided Channel State Information," *submitted to IEEE Journal on Selected Topics in Signal Processing*, 2009.

[55] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[56] H. Bolcskei, "Principles of MIMO-OFDM Wireless Systems," *Signal Processing for Mobile Communications Handbook, CRC Press*, 2004.

[57] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C (2nd ed.): the art of scientific computing*. New York, NY, USA: Cambridge University Press, 1992.

[58] J.-J. Chang, D.-J. Hwang, and M.-C. Lin, "Some extended results on the search for good convolutional codes," *IEEE Transactions on Information Theory*, vol. 43, no. 5, pp. 1682–1697, September 1997.

[59] J. Hansen and H. Bolcskei, "A geometrical investigation of the rank-1 Ricean MIMO channel at high SNR," *IEEE International Symposium on Information Theory 2004*, p. 64, 27 June-2 July 2004.

[60] H. Zheng and H. Viswanathan, "Optimizing the ARQ performance in downlink packet data systems with scheduling," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 495–506, March 2005.

[61] H. Bolcskei, M. Borgmann, and A. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 427–439, April 2003.

[62] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, October 1998.

[63] L. Fang and L. Milstein, "Successive interference cancellation in multicarrier DS/CDMA," *IEEE Transactions on Communications*, vol. 48, no. 9, pp. 1530–1540, September 2000.